THESIS REPORT
Ph.D.

Nonholonomic Geometry, Mechanics and Control

by R. Yang
Advisor: P.S. Krishnaprasad

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **1992**	2. REPORT TYPE	3. DATES COVERED **00-00-1992 to 00-00-1992**

4. TITLE AND SUBTITLE **Nonholonomic Geometry, Mechanics and Control**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland,The Graduate School,2123 Lee Building,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **180**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

NONHOLONOMIC GEOMETRY, MECHANICS AND CONTROL

By

Rui Yang

Advisory Committee:

Professor P. S. Krishnaprasad, Chairman/Advisor
Professor C. A. Berenstein
Professor S. I. Marcus
Associate Professor M. A. Shayman
Associate Professor W. P. Dayawansa

ABSTRACT

Title of Dissertation: Nonholonomic Geometry, Mechanics and Control

Rui Yang, Doctor of Philosophy, 1992

Dissertation directed by: P. S. Krishnaprasad, Professor

Department of Electrical Engineering

This dissertation is concerned with dynamic modeling and kinematic control of constrained mechanical systems with symmetry from a geometric point of view. Constraints are defined via the characteristics of distributions or codistributions on the tangent bundle (velocity phase space) of configuration space. Lie symmetry groups acting on the systems are assumed to leave both Lagrangian and constraints invariant. As a special case of mechanical systems with holonomic constraints, we rigorously analyze the kinematics and dynamics of floating, planar four-bar linkages. The analyses include topological description of the configuration space, symplectic and Poisson reductions of the dynamics and bifurcation of relative equilibria. For kinematic control of nonholonomic systems, we mainly study the related optimal control problem for a system consisting of a rigid body with two oscillators. In particular, the intrinsic formulation and explicit solvability of necessary conditions for the optimal control are investigated from a Hamiltonian point of view. In the study of the dynamics of Lagrangian systems with constraints, the nonholonomic distributions are defined via arbitrary choices of principal connections. We show that, under our hypotheses on constraints and exterior force, the dynamics of a nonholonomic Lagrangian system with non-Abelian symmetry can be reduced to a lower dimensional space determined by the principal fiber bundle. The reduced dynamic equations are formulated explicitly. This formulation generalizes the one for classical Chaplygin systems which possess Abelian symmetry, and the one having non-Abelian symmetry but with linear constraints. In addition, if a special principal connection, that is, the mechanical connection by Kummer and Smale, is considered, our formulation for nonholonomic systems also leads to the one in Lagrangian reduction discovered recently by Marsden and Scheurle. The results of this dissertation have direct application in space robotics and nonholonomic motion planning in robotics.

DEDICATION

To

My Parents and Sister

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The motions of various mechanical systems which we wish to synthesize and control often have to satisfy certain kinds of restrictions imposed by the natural environment or the structure of the systems themselves. In mechanics, such restrictions are called *constraints*. Although the fundamental theory of mechanical systems with constraints was established and developed in the last century, recent research and developments in analytical mechanics and control theory from a geometric viewpoint have inspired a strong desire to reinterpret and reformulate the theory of constrained dynamics in an intrinsic geometric way. In addition, many practical problems in recent investigations in mechanical and electrical engineering, such as modeling and control of mobile robots and dextrous robotic hands, and the design and control of spacecraft, also show the need for a deeper understanding of the role that constraints play in mechanical systems.

In classical Lagrangian mechanics, constraints are usually classified into two types: holonomic and nonholonomic constraints. Roughly speaking, if we exclude those constraints which are given by a set of inequalities or time dependent equations, the former imposes a limitation on the motion of the system to a subspace of the configuration space, but the latter imposes a limitation on the motion of the system to a subspace of

1

the velocity phase space. If we further restrict our attention to constraints which can be expressed via a set of linear functions of velocities of the motion, using Frobenius' theorem in differential geometry, the distinction between two kinds of classical constraints can be given rigorously via integrability of corresponding distributions. From this point of view, a holonomic constraint is characterized by an integrable distribution, and the motion of the system under study is restricted to a leaf of the corresponding foliation. Then, using suitable local coordinates, one can always describe the motion of the constrained system by a set of unconstrained dynamic equations on a lower dimensional space. Because of this, in analytical mechanics, establishing the dynamic equations and determining the motion for systems with holonomic constraints are not considered as challenging problems in general. However, from a practical point view, it is always of interest to understand how the constraints affect the complexity of dynamic motions of a specific system. In particular, such understanding usually generates sophisticated strategies for design and control of the system. On the other hand, using geometric interpretations, classical nonholonomic constraints can be characterized by nonintegrable distributions on the configuration space. Then the real motion of a mechanical system with nonholonomic constraints must satisfy the condition that the tangent vector of a path in configuration space belongs to such distribution.

Treating the classical constraints in the geometric way as above, one can think of more general questions related to the theory of distribution on manifolds and various nonholonomic problems. For instance, it is known that the motion of a Lagrangian mechanical system can be described completely by a *special* vector field on the tangent bundle of its configuration space (i.e., via second order differential equations). Therefore, from a geometric point of view, it is more natural to consider the constraints in terms of distributions on the tangent bundle. In doing so, certain constraints which are not linear in velocity but, for example, those described by the null set of some set of functions on tangent bundle, can also be given geometric interpretations. Hence, in a certain sense, a geometric treatment generalizes the range of constrained problems we can deal with. A natural question that follows is that of reformulating the Lagrange-

d'Alembert principle in an intrinsic form. In [46,47], this and related issues are carefully studied. There is another viewpoint in seeking a generalization of constrained problems in mechanics. Notice that the concept of distribution in differential geometry is concerned only with the mathematical meaning of a constraint. In dealing with mechanics, classical constraints (e.g., geometric and kinematic constraints) are just one means to define or construct related distributions. This suggests that one can think of other kinds of constraints coming from, for instance, conservation laws as a source of distributions. In [5], the former are called phenomenological constraints, the latter are called symmetry constraints.

In mathematics, general nonholonomic problems are those of determining a class of curves in terms of a certain law or principle or criterion such that it satisfies restrictions given by distributions on a given space. From this point of view, it is clear that the curves determined from constrained classical mechanical systems satisfy the Lagrange-d'Alembert principle. One can certainly consider such problems in terms of other criteria or principle, for instance, Hamilton's principle of least action. Of course, it has been known that the latter does not give real motions for classical mechanical systems if the constraints are nonholonomic. Recent research in this direction leads to certain nonstandard problems of geodesics [48]. In [8] these are referred to as problems of singular Riemannian geometry and Strichartz speaks of sub-Riemannian geometry [45]. In mechanics, this problem is closely related to the problem of nonholonomic motion planning or kinematic control [22,23,31].

This dissertation is concerned with both types of problems with constraints mentioned above. The main feature here is that we focus on systems which possess group symmetries. The fundamental mathematical tools applied here are those which have been widely used in geometric mechanics, such as Lie groups and Lie algebras, symplectic and Poisson geometry, Lagrangian and Hamiltonian mechanics, reductions and connection theory on principal fiber bundles. This dissertation is organized as follows.

In Chapter II, after reviewing some definitions, notations and important theorems in differential geometry and geometric mechanics, we introduce an invariant/intrinsic

formulation of Lagrangian mechanics due to Vershik [46] and Vershik and Faddeev [47]. Such a formulation makes it possible for us to restate Lagrange-d'Alembert principle with constraints on second tangent bundle using the concept of virtual displacement in classical mechanics. It is easy to show that this formulation is equivalent to the one using constraint reaction force considered in [46,47]. But, our expression of Lagrange-d'Alembert principle for constrained systems turns out to be very important in Chapter V, in which symmetries are associated to the systems under study.

In Chapter III, we study the kinematics and dynamics of the simplest coupled mechanical systems with holonomic constraints, that is the planar, *floating* four-bar linkages. Although a floating four-bar linkage is a particular mechanism with holonomic constraint, it illustrates precisely what we mentioned earlier, that is, its dynamic properties in comparison with an open chain are very different. This chapter provides complete analyses including topological description of the configuration space, symplectic and Poisson reductions of the dynamics and bifurcation of relative equilibria.

Chapter IV addresses reduction and explicit solvability of optimal control problems on principal bundles with connections from a Hamiltonian point of view. The particular mechanical system we consider is a rigid body with two oscillators. The optimal control problem is posed by considering a special nonholonomic variational problem, in which the nonholonomic distribution is defined via a connection. The necessary conditions for the optimal control problem are determined intrinsically by a perturbation method and a Hamiltonian formulation. The necessary conditions admit the structure group of the principal bundle as a symmetry group of the system. Thus the problem is amenable to Poisson reduction. Under suitable hypotheses and approximations, we find that the reduced system possesses additional Abelian symmetry. Applying Poisson reduction again, we obtain a further reduced system and corresponding first integral. The model problem described here is strongly motivated by a troublesome phenomenon of drift observed in the Hubble Space Telescope due to thermo-elastically driven vibrations of the solar panels arising from the day-night thermal cycling on-orbit. The point mass oscillators in our problem may be viewed as one-mode truncations of the elasto-

4

mechanical problem.

The most important contribution of this dissertation to nonholonomic Lagrangian mechanics is contained in Chapter V, in which we discover a *reduction theorem* for the Lagrangian systems with non-Abelian symmetries and certain nonholonomic constraints. This chapter is motivated by Chaplygin's reduction formulation in classical mechanics with Abelian symmetry. It is known that, in general for a system with nonholonomic constraints, one cannot get reduced dynamics on a lower dimensional space by eliminating constraints, as one can do for holonomic constraints. Instead, one usually has to expand the space by bringing in more variables, i.e., Lagrange multipliers. But Chaplygin showed that one can do such reduction if the system and constraints admit an Abelian symmetry. A natural question one can ask is if this is also possible for constrained systems which admit a non-Abelian symmetry? This question is answered in some detail in this chapter. The constraints here are constructed on principal fiber bundles with connections. Under our assumption, they take the form of affine functions in velocity, instead of linear ones. In using distribution theory to interpret constraints, one has to consider the distributions on the tangent bundle of the configuration space, rather than on configuration space. The approach of constructing constrained dynamics on the second tangent bundle introduced in Chapter II can be directly applied here. Since we consider constraints from principal connections which decompose the velocity phase space, the dynamics of the system can be described on a horizontal subspace and, consequently, on the tangent bundle of the quotient space of the configuration space with respect to the symmetry group. Then our goal to reduce the dynamics on a lower dimensional space can be reached. Following this idea, we obtain a reduction theorem. An important application of our reduction theorem is the derivation of Lagrangian reduction, where the principal connection used is the mechanical connection determined by the conserved momentum map. In this application, our result coincides with the one in [27]. Some representative examples are also studied at the end of this chapter.

In Chapter VI, we reconsider the main results in this dissertation and point out some topics for future research.

CHAPTER II

PRELIMINARIES

In this chapter, we first review some definitions, notations and important theorems in differential geometry and geometric mechanics. Concepts and results that will be used frequently in the following chapters are the main focus of this chapter. The second aim of this chapter is to introduce an invariant formulation of Lagrangian mechanics due to Vershik [46] and Vershik and Faddeev [47]. In particular, we are interested in the statement of Lagrange-d'Alembert principle for systems with constraints. An important feature of Section 2.2 is that we work with constraints on vector fields that behave as virtual displacements in classical mechanics. Such a presentation leads to the form of Lagrange-d'Alembert principle expressed in terms of reaction forces due to constraints, which is the form given in [46,47]. Our expression of Lagrange-d'Alembert principle for constrained systems will play a crucial role in Chapter V, where symmetries will be associated to the systems under study.

2.1 Geometric Mechanics

In this section, we recall some basic definitions, commonly used notations and important theorems in geometric mechanics, which will be cited in the following chapters. The useful references are [1,2].

2.1.1 Lie Groups and Group Actions

A *Lie group* G is a differentiable manifold and a group, for which the group operations, product and inverse, are differentiable maps. We denote by $R_g : G \to G : h \mapsto hg$ and $L_g : G \to G : h \mapsto gh$ right and left translation, respectively. One can show that the tangent space of G at identity e, $T_e G$, forms a Lie algebra which is isomorphic to the set of left invariant vector fields on G (see [1], Chapter 4). We denote by $\mathcal{G}$ the Lie algebra of G. The Lie bracket of ξ and η is denoted as $[\,\xi, \eta\,], \forall \xi, \eta \in \mathcal{G}$.

Let Q be a smooth manifold. A *(left) action* of a Lie group G on Q is a smooth mapping

$$\Phi : G \times Q \to Q : \quad (g, q) \mapsto \Phi(g, q) \overset{\triangle}{=} \Phi_g(q) \overset{\triangle}{=} g \cdot q \tag{2.1.1}$$

such that $\Phi(e, q) = q, \forall q \in Q$, and $\Phi(g, \Phi(h, q)) = \Phi(gh, q), \forall g, h \in G, \forall q \in Q$. Certain induced actions of G on manifolds M are of particular interest. Let $M = TQ$, the tangent bundle of Q. The *tangent lift action* of G on TQ is given by

$$\Phi^T : G \times TQ \to TQ : \quad (g, v_q) \mapsto T_q \Phi_g v_q, \tag{2.1.2}$$

where $T_q \Phi_q$ denotes the linearization of Φ_g at $q \in Q$. Let $M = T^*Q$, the cotangent bundle of Q. The *cotangent lift action* of G on T^*Q is given by

$$\Phi^{T^*} : G \times T^*Q \to T^*Q : \quad (g, \alpha_q) \mapsto T_q^* \Phi_{g^{-1}} \alpha_q, \tag{2.1.3}$$

where $T_q^* \Phi_g$ is the linear dual of $T_q \Phi_q$. Let $M = \mathcal{G}$. The *adjoint action* of G on $\mathcal{G}$ is defined by

$$Ad : G \times \mathcal{G} \to \mathcal{G} : \quad (g, \xi) \mapsto Ad_g \xi \overset{\triangle}{=} T_e(R_{g^{-1}} L_g) \xi. \tag{2.1.4}$$

Let $M = \mathcal{G}^*$, the dual of Lie algebra $\mathcal{G}$ of G. Then the *co-adjoint action* of G on $\mathcal{G}^*$ is defined by

$$Ad^* : G \times \mathcal{G}^* \to \mathcal{G}^* : \quad (g, \nu) \mapsto Ad_{g^{-1}}^* \nu, \tag{2.1.5}$$

where the operation Ad_g^* is given by $\langle Ad_g^* \nu, \eta \rangle = \langle \nu, Ad_g \eta \rangle$, for any $\eta \in \mathcal{G}$.

Let $exp(t\xi)$ be the integral curve of the left invariant vector field on G associated to $\xi \in \mathcal{G}$. For a group action as in (2.1.1) and $\xi \in \mathcal{G}$, the vector field defined by

$$\xi_Q(q) = \left.\frac{d}{dt}\right|_{t=0} \Phi(exp(t\xi), q), \quad \forall q \in Q \tag{2.1.6}$$

is called *infinitesimal generator* or *fundamental vector field* of the action corresponding to ξ. One can show that if the group G acts on Q freely and effectively, the mapping, $\xi \mapsto \xi_Q$ is an isomorphism of the Lie algebra $\mathcal{G}$ into the Lie algebra $\mathcal{X}(Q)$, the set of all smooth vector fields on Q (cf. [33]). For the adjoint action defined as in (2.1.4),

$$\xi_{\mathcal{G}}(\eta) = ad_\xi \eta \overset{\triangle}{=} [\xi, \eta] \quad \forall \eta \in \mathcal{G}. \tag{2.1.7}$$

For the co-adjoint action defined in (2.1.5),

$$\xi_{\mathcal{G}^*}(\alpha) = -ad_\xi^* \alpha, \quad \forall \alpha \in \mathcal{G}^*. \tag{2.1.8}$$

2.1.2 Hamiltonian Systems

We first consider Hamiltonian systems on a symplectic manifold. A symplectic manifold (M, Ω) is an even dimensional, smooth manifold M together with a closed, nondegenerate two-form Ω, called *symplectic structure*, on M. Given a smooth function H, called the *Hamiltonian or energy function*, on M, the *Hamiltonian vector field* X_H is the smooth vector field on M satisfying

$$\Omega(X_H, Y) = dH \cdot Y \quad \forall Y \in \mathcal{X}(M). \tag{2.1.9}$$

Let Q be an n-dimensional smooth manifold, which represents the configuration space of a mechanical system. Let $M = T^*Q$ be the momentum phase space, coordinatized locally by $z = (q_1, \cdots, q_n, p_1, \cdots, p_n)$. One can show that a symplectic structure on T^*Q is a *canonical* two-form which is represented locally by $\Omega_0 = \sum_{i=1}^n dq_i \wedge dp_i$ (cf. Theorem 3.2.10 in [1]). Then the Hamiltonian vector field is given locally by $X_H = \Lambda \nabla H$, where matrix $\Lambda = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$.

One can also work with a symplectic structure on $M = TQ$, the velocity phase space. Given a smooth function L, called *Lagrangian*, on TQ, the *Legendre transformation* is defined by fiber derivative $\mathbf{F}L : TQ \to T^*Q$ associated to L. Let $\Omega_L \triangleq (\mathbf{F}L)^*\Omega_0$. One can show that if L is regular, Ω_L is symplectic. Hence, (TQ, Ω_L) is a symplectic manifold. Let $(q, v) \in TQ$ and the energy function

$$H_L(q, v) \triangleq \mathbf{F}L(v) \cdot v - L(q, v).$$

A vector field X_{H_L}, called *Lagrangian vector field*, on TQ is determined by

$$\Omega(X_{H_L}, Y) = dH_L \cdot Y, \quad \forall Y \in \mathfrak{X}(TQ). \tag{2.1.10}$$

One shows that if L is hyperregular, i.e., $\mathbf{F}L$ is a diffeomorphism and the Hamiltonian $H \triangleq H_L \circ (\mathbf{F}L)^{-1}$, then the base integral curves given by bundle projections of the integral curves of X_{H_L} and X_H on Q are identical.

We next consider Hamiltonian systems defined on Poisson manifolds. A Poisson manifold $(M, \{\cdot, \cdot\})$ is a smooth manifold M together with an $\Re$-bilinear map on $C^\infty(M)$:

$$\{\cdot, \cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M),$$

referred to as *Poisson structure* or *Poisson bracket*, which satisfies the following axioms: for $f_i \in C^\infty(M), i = 1, 2, 3$,

1) Skew Symmetry: $\{f_1, f_2\} = -\{f_2, f_1\}$;

2) Leibniz' Rule: $\{f_1, f_2 f_3\} = \{f_1, f_2\} f_3 + f_2 \{f_1, f_3\}$;

3) Jacobi Identity: $\{f_1, \{f_2, f_3\}\} + \{f_3, \{f_1, f_2\}\} + \{f_2, \{f_3, f_1\}\} = 0$.

A Poisson structure can be expressed uniquely through a contravariant skew-symmetric two-tensor Λ on M, called *Poisson tensor*, such that

$$\{f, g\}(z) = \Lambda(z)(df(z), dg(z)), \qquad \forall z \in M. \tag{2.1.11}$$

If M is an n-dimensional manifold (n is finite), Λ is given by an $n \times n$ skew-symmetric matrix, also denoted as Λ, and the Poisson bracket can be expressed as

$$\{f, g\}(z) = \nabla f^T(z) \Lambda(z) \nabla g(z). \tag{2.1.12}$$

Given a smooth function H, referred to as *Hamiltonian*, on M, the *Hamiltonian vector field* on $(M, \{\cdot, \cdot\})$ is defined by the relation

$$X_H(f) = \{f, H\}. \tag{2.1.13}$$

The uniqueness of such a vector field can be proved, cf. [30]. Let $\phi(t)$ be the flow of X_H and Z be a coordinate function, we then have

$$\frac{d}{dt}(Z(\phi(t))) = \{Z(\phi(t)), H\}. \tag{2.1.14}$$

In the finite dimensional case, we have

$$\dot{z} = \Lambda(z)\nabla H(z), \tag{2.1.15}$$

where $z(t) = Z(\phi(t))$. A function C on M is a *Casimir function* if $\{C, F\} = 0$ for all $F \in C^\infty(M)$, i.e., C is a constant along the flow of *all* Hamiltonian vector fields.

Remark 2.1.1: Every symplectic manifold is a Poisson manifold by defining the Poisson structure to be $\{f, g\} = \Omega(X_f, X_g)$. It is a theorem of Lie and Kirillov that, in general, Poisson manifolds can be decomposed into the symplectic leaves of a generalized foliation [25]. As an example, the Lie-Berezin-Kirillov-Kostant-Souriau Poisson structures on the dual space of Lie algebra $\mathcal{G}$, $\mathcal{G}^*$, is considered. One can show that the symplectic leaf through each point μ in $\mathcal{G}^*$ is the coadjoint orbit through μ (cf. [21]). ∎

2.1.3 Symmetry and Reduction

Let $(M, \{ \, , \, \})$ be a Poisson manifold and let Lie group G act on $(M, \{ \, , \, \})$ canonically, that is, for each $g \in G$, the map $\Phi_g : M \to M$ preserves the Poisson structure, i.e.,

$$\{f_1, f_2\} \circ \Phi_g = \{f_1 \circ \Phi_g, f_2 \circ \Phi_g\}, \forall f_i \in C^\infty(M). \tag{2.1.16}$$

A *momentum mapping*, $\mathbf{J} : M \to \mathcal{G}^*$ of the action is defined by

$$\langle \mathbf{J}(z), \xi \, \rangle = J(\xi)(z), \tag{2.1.17}$$

10

for all $\xi \in \mathcal{G}$ and $z \in M$, where a linear map $J : \mathcal{G} \to C^\infty(M)$ is assumed to exist such that $X_{J(\xi)} = \xi_M$, or from (2.1.13),

$$\{f, J(\xi)\} = \xi_M(f), \qquad \forall f \in C^\infty(M). \tag{2.1.18}$$

A momentum map $\mathbf{J}$ is called Ad^*-equivariant if $\mathbf{J} \circ \Phi_g = Ad^*_{g^{-1}}\mathbf{J}$, where Ad^* is the co-adjoint action defined in (2.1.15).

It is well-known that, if the Hamiltonian H is G-invariant, the momentum map provides a first integral for the system. This result is known as Noether's theorem (in Hamiltonian version) which is stated below.

Theorem 2.1.2: If the Lie group, G, acting canonically on the Poisson manifold M admits a momentum mapping $\mathbf{J} : M \to \mathcal{G}^*$ and $H \in C^\infty(M)$ is G-invariant, i.e., $H \circ \Phi_g = H$ or $\xi_M(H) = 0$ for all $\xi \in \mathcal{G}$, then $\mathbf{J}$ is a constant of the motion for H, i.e., $\mathbf{J} \circ \phi_t = \mathbf{J}$, where ϕ_t is the flow of X_H. ∎

Next, we introduce the general framework of *Poisson reduction*. Let a Lie group G acts on Poisson manifold $(M, \{\ ,\ \}_M)$ freely and effectively. Let $\pi : M \to M/G$ be the canonical projection. Here we assume the quotient space M/G is a smooth manifold. $(M, \{\ ,\ \}_M)$ is *Poisson reducible* if M/G has a Poisson structure $\{\ ,\ \}_{M/G}$ which satisfies

$$\{f \circ \pi, h \circ \pi\}_M = \{f, h\}_{M/G} \circ \pi, \tag{2.1.19}$$

for $f, h \in C^\infty(M/G)$. If the Hamiltonian H on M is G-invariant, the reduced Hamiltonian $\tilde{H}$ on M/G is defined by

$$\tilde{H} \circ \pi = H. \tag{2.1.20}$$

In addition, the Poisson reduced Hamiltonian vector field, $X_{\tilde{H}}$ is given by

$$X_{\tilde{H}}(\tilde{f}) = \{\tilde{f}, \tilde{H}\}_{M/G}, \qquad \forall \tilde{f} \in C^\infty(M/G). \tag{2.1.21}$$

Remark 2.1.3: Another point of view of reduction in geometric mechanics is the *symplectic reduction*. For the detail of its framework, cf. [1,28], and for an application, cf. [43]. ∎

2.1.4 Simple Mechanical Systems with Symmetry

A *simple mechanical system with symmetry* is a 4-tuple (Q, K, V, G), where,

(i) Q is a smooth manifold, the configuration space of the system;

(ii) K is Riemannian metric on Q, whose value on $T_q Q \times T_q Q$, for each $q \in Q$, is written as $K(q)(v_q, w_q)$, $\forall v_q, w_q \in T_q Q$. The kinetic energy of the system is then written as $\frac{1}{2} K(q)(v_q, v_q)$;

(iii) V is a function on Q, the potential energy;

(iv) G is a connected Lie group with an action $\Phi : G \times Q \to Q$ which leaves both the Riemannian metric and the potential energy invariant, i.e., for each q in Q and g in G,

$$K(\Phi_g(q))(T_q \Phi_q \cdot v_q, T_q \Phi_q \cdot w_q) = K(q)(v_q, w_q), \quad \forall v_q, w_q \in T_q Q \qquad (2.1.22)$$

and $V(\Phi_g(q)) = V(q)$.

For a given simple mechanical system with symmetry (Q, K, V, G), the associated Lagrangian is defined by

$$L : TQ \to \Re : \quad v_q \mapsto L(v_q) = \frac{1}{2} K(q)(v_q, v_q) - V \circ \tau(v_q), \qquad (2.1.23)$$

where $\tau : TQ \to Q$ is the canonical tangent projection. The Legendre transform $\mathbf{F}L$ of L is given here by the vector bundle isomorphism

$$K^\flat : TQ \to T^*Q \qquad (2.1.24)$$

satisfying $K^\flat(v_q) \cdot w_q = K(q)(v_q, w_q), \forall v_q, w_q \in T_q Q$. The Hamiltonian $H : T^*Q \to \Re$ is defined to be

$$H(\alpha_q) = \frac{1}{2} K(q)((K^\flat)^{-1}(\alpha_q), (K^\flat)^{-1}(\alpha_q)) + V \circ \tau^*(\alpha_q), \qquad (2.1.25)$$

where $\tau^* : T^*Q \to Q$ is the canonical cotangent projection. The standard momentum map can be defined either by $\mathbf{J} : T^*Q \to \mathcal{G}^*$ for the lifted action Φ^{T^*} given in (2.1.3) such that, for each $\xi \in \mathcal{G}$ and $\alpha_q \in T_q^*Q$,

$$\langle \mathbf{J}(\alpha_q), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle, \qquad (2.1.26)$$

or by $\mathbf{J}^{\sharp} : TQ \to \mathcal{G}^*$ for the lifted action Φ^T given in (2.1.4) such that, for each $\xi \in \mathcal{G}$ and $v_q \in T_q Q$,

$$\langle \mathbf{J}^{\sharp}(v_q), \xi \rangle = K(q)(v_q, \xi_Q(q)). \tag{2.1.27}$$

One can show that both momentum maps are Ad^*-equivariant. (In the above equations and later, we use $\langle \cdot, \cdot \rangle$ to denote the natural pairing of the elements in a vector space and its dual. We sometimes also use "$\cdot$" to indicate this pairing.)

Remark 2.1.4: Addition of linear in velocity term to the Lagrangian of a simple mechanical system with symmetry gives rise to gyroscopic systems with symmetry. For a careful study of such systems, see [49]. ■

2.2 Lagrangian Mechanics with Constraints

In this section we present Lagrange-d'Alembert principle for Lagrangian systems with constraints in invariant form. We first recall some important definitions and facts. The proofs for most of those facts can be found in [46,47] and will be omitted here. Although our treatment follows these references, one should notice the differences of certain sign conventions. In [50] the intrinsic Lagrangian viewpoint is exploited in the study of gyroscopic stabilization.

Let Q be a smooth, n-dimensional manifold and TQ the tangent bundle of Q with the canonical projection $\pi : TQ \to Q$. Let TTQ be the second tangent bundle of Q and $T\pi : TTQ \to TQ$ be the tangent map of π. In local coordinates, one has $\pi : (q, v) \mapsto q$ at each point (q, v) in TQ and

$$\begin{aligned} T\pi_{(q,v)} : T_{(q,v)} TQ &\to T_q Q \\ (u, w) &\mapsto (q, u), \end{aligned} \tag{2.2.1}$$

where $u, w \in T_q Q$.

Let $\mathfrak{X}(TQ)$ be the set of vector fields on TQ. A vector field, $X \in \mathfrak{X}(TQ)$, is *special* if $T\pi \circ X$ is identity. In local coordinates, at each point $(q, v) \in TQ$,

$$X(q, v) = (v, w), \tag{2.2.2}$$

for some $w \in T_q Q$. It is clear that a special vector field defines a *differential equation of second order*.

Let $T_{(q,v)}^V TQ$ be a subspace of $T_{(q,v)}TQ$ such that each vector in $T_{(q,v)}^V TQ$ is tangent to the fiber of TQ. In local coordinates, every vector in $T_{(q,v)}^V TQ$ has the form of $(0, w)$ for some $w \in T_q Q$. This subspace is referred to as *vertical* tangent subspace. It follows that one can identify $T_q Q$ and $T_{(q,v)}^V TQ$ by isomorphism through the mapping

$$\gamma_{(q,v)} : T_q Q \hookrightarrow T_{(q,v)}TQ$$
$$w \mapsto (0, w). \tag{2.2.3}$$

A vector field $X^V \in \mathfrak{X}(TQ)$ is *vertical* if, at each point (q, v) in TQ, $X^V(q, v) \in T_{(q,v)}^V TQ$. A vector field $X^{pr} \in \mathfrak{X}(TQ)$ is *principal* if it is vertical and, at each point (q, v) in TQ, $X^{pr}(q, v) = \gamma_{(q,v)} \cdot v$, or explicitly, $X^{pr}(q, v) = (0, v)$. Thus, there is a unique such vector field.

We next consider the corresponding geometric objects and maps on the dual spaces. Let T^*Q and T^*TQ be the cotangent bundle of Q and TQ, respectively. Let $\varpi^1(TQ)$ be the totality of one-forms on TQ. A 1-form $\omega \in \varpi^1(TQ)$ is *horizontal* if it annihilates vertical vector fields on TQ. In local coordinates, $(\alpha, \beta) \in T_{(q,v)}^* TQ$ for $\alpha, \beta \in T_q^* Q$ is horizontal if and only if $\beta = 0$. Let $(T\pi)_{(q,v)}^* : T_q^* Q \to T_{(q,v)}^* TQ$ be the linear dual of the map $T\pi_{(q,v)}$ given in (2.2.1). If $\alpha \in T_q^* Q$, one can show that

$$(T\pi)_{(q,v)}^*(\alpha) = (\alpha, 0), \tag{2.2.4}$$

which is horizontal. Indeed, if $(u, w) \in T_{(q,v)}TQ$,

$$\langle (T\pi)_{(q,v)}^*(\alpha), (u, w) \rangle = \langle \alpha, (T\pi)_{(q,v)}(u, w) \rangle = \langle \alpha, u \rangle = \langle (\alpha, 0), (u, w) \rangle.$$

Moreover, let $\gamma_{(q,v)}^* : T_{(q,v)}^* TQ \to T_q^* Q$ be the dual of the map $\gamma_{(q,v)}$ given in (2.2.3). One can show that

$$\gamma_{(q,v)}^*(\alpha, \beta) = \beta, \tag{2.2.5}$$

where $\alpha, \beta \in T_q^* Q$. Indeed, $\langle \gamma_{(q,v)}^*(\alpha, \beta), u \rangle = \langle (\alpha, \beta), \gamma_{(q,v)}u \rangle = \langle (\alpha, \beta), (0, u) \rangle = \langle \beta, u \rangle$.

Define a bundle map $\tau : T^*TQ \to T^*TQ$ by

$$\tau \overset{\triangle}{=} (T\pi)^* \gamma^*. \tag{2.2.6}$$

From (2.2.4) and (2.2.5) one shows that, for $(\alpha, \beta) \in T_{(q,v)}^* TQ$,

$$\tau_{(q,v)}(\alpha, \beta) = (\beta, 0), \tag{2.2.6'}$$

14

which is horizontal. (Indeed, $\tau_{(q,v)}(\alpha,\beta) = (T\pi)^*\gamma^*_{(q,v)}(\alpha,\beta) = (T\pi)^*\beta = (\beta,0)$.)

Remark 2.2.1: It should be kept in mind that the horizontal/vertical operations defined in this section are based on the structure of the second tangent bundle of a manifold. They are different from the horizontal/vertical operations on the principal fiber bundle, which will be introduced in Chapter IV. ∎

We are now ready to define some geometric objects to present *Lagrange-d'Alembert principle.* Let the smooth manifold Q be the configuration space of a mechanical system and let a smooth function L on TQ be the *Lagrangian.* Define a horizontal 1-form by

$$\omega_L \overset{\triangle}{=} \tau \circ dL, \tag{2.2.7}$$

where $dL : TQ \to T^*TQ$ is exterior derivative of L. In local coordinates, $dL(q,v) = (D_1 L(q,v), D_2 L(q,v))$, where $D_i L(q,v)$ is the Fréchét derivative of L relative to i-th argument of L. Then,

$$\omega_L(q,v) = (D_2 L(q,v), 0). \tag{2.2.7$'$}$$

Define a 2-form, Ω_L, by

$$\Omega_L \overset{\triangle}{=} -d\omega_L, \tag{2.2.8}$$

where $d\omega_L$ is the exterior derivative of ω_L. From (2.2.7)$'$ it is easy to obtain the expression for Ω_L in local coordinates, that is,

$$\Omega_L(q,v)((u_1,w_1),(u_2,w_2)) = (D_1 D_2 L(q,v) \cdot u_2) \cdot u_1 + (D_2 D_2 L(q,v) \cdot w_2) \cdot u_1$$
$$- (D_1 D_2 L(q,v) \cdot u_1) \cdot u_2 - (D_2 D_2 L(q,v) \cdot w_1) \cdot u_2, \tag{2.2.8$'$}$$

where $(u_i, w_i) \in T_{(q,v)}TQ, i = 1, 2$. Define the energy function H_L on TQ by

$$H_L \overset{\triangle}{=} dL(X^{pr}) - L, \tag{2.2.9}$$

which, in local coordinates, can be written as

$$H_L(q,v) = D_2 L(q,v) \cdot v - L(q,v). \tag{2.2.9$'$}$$

Finally, define *Lagrangian force* on a special vector field X, denoted as $F_L(X)$, by

$$F_L(X) \overset{\triangle}{=} \Omega_L(X, \cdot) - dH_L. \tag{2.2.10}$$

One can show that $F_L(X) \in \varpi^1(TQ)$ is horizontal and, from $(2.2.8)'$ and $(2.2.9)'$, its expression in local coordinates is, for $X(q,v) = (v,w)$,

$$F_L(X)(q,v)(u_1,u_2) = (-D_1 D_2 L(q,v) \cdot v - D_2 D_2 L(q,v) \cdot w + D_1 L(q,v)) \cdot u_1, \quad (2.2.10)'$$

for any $(u_1, u_2) \in T_{(q,v)}TQ$.

Using the above notions, the Lagrange-d'Alembert principle for Lagrangian system without constraints is given as follows.

Lagrange-d'Alembert Principle (*without Constraints*) **2.2.2:**

On the special vector field which determines the real trajectory of motion, the sum of the Lagrangian force and exterior force vanishes on *any* vector field on the tangent bundle of configuration space. ∎

More explicitly, this principle says that the special vector field X whose integral curves are the trajectories of motion in TQ satisfies

$$\langle F_L(X) + \omega_e, Z \rangle = 0, \quad (2.2.11)$$

for *any* vector field $Z \in \mathfrak{X}(TQ)$, where ω_e is a horizontal one-form on TQ denoting the exterior force. Following $(2.2.10)'$, the local expression of $(2.2.11)$ can be given as, at each point $(q,v) \in TQ$ with $Z(q,v) \stackrel{\triangle}{=} (u,w)$ and $\omega_e(q,v) \stackrel{\triangle}{=} (\alpha_e, 0)$,

$$(-D_1 D_2 L(q,v) \cdot v - D_2 D_2 L(q,v) \cdot w + D_1 L(q,v) + \alpha_e) \cdot u = 0$$

or, by letting $v = \dot{q}$ and $w = \dot{v}$,

$$(\frac{d}{dt} D_2 L(q,v) - D_1 L(q,v)) \cdot u = \alpha_e \cdot u. \quad (2.2.12)$$

For finite dimensional systems, $(2.2.12)$ gives the classical Euler-Lagrange equation, (in vector form)

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \alpha_e. \quad (2.2.13)$$

We next introduce an important mapping and two lemmas which will be used later. If the Lagrangian L is regular, one shows that Ω_L is nondegenerate. This gives rise to a well defined one-to-one mapping, $\Pi_L : \varpi^1(TQ) \rightarrow \mathfrak{X}(TQ)$, determined by

$$\Omega_L(\Pi_L(\omega), Y) = \omega(Y), \quad (2.2.14)$$

16

for any one-form ω and any vector field Y on TQ. The following lemma gives an important property of this mapping [47].

Lemma 2.2.3: Let L be regular. If ω is horizontal one-form on TQ and $\omega \neq 0$, then $\Pi_L(\omega) \neq 0$ and $\Pi_L(\omega)$ is a vertical vector field on TQ. ∎

Define the *Lagrangian vector field*, X_{H_L}, by

$$X_{H_L} \triangleq \Pi_L(dH_L), \tag{2.2.15}$$

where H_L is the energy function given in (2.2.9). The following useful lemma is also shown in [47].

Lemma 2.2.4: X_{H_L} is a special vector field on TQ. ∎

Before we proceed to the discussion of constrained systems, we note:

Remark 2.2.5: One can show that the geometric objects Ω_L, H_L and X_{H_L} here are exactly the same as those in the Section 2.2.1. ∎

Remark 2.2.6: It is very important to realize that (2.2.11) means that the *work* done by the sum of Lagrangian force and exterior force on Z is zero. This work is, in fact, *virtual work* and Z is *virtual displacement* (on TQ) in the literature of classical mechanics. Therefore (2.2.11) is nothing but the principle of virtual work. It is also important to note that when the motion of the system is under certain constraints (holonomic or nonholonomic), this principle is also true provided that the vector field Z satisfies further conditions relating to the constraints. This is what we will consider next. ∎

Before defining constraints, we first give our definitions of a distribution and its annihilator.

Definition 2.2.7: Let M be a smooth manifold and TM be its tangent bundle with projection $pr : TM \to M$. Let $M' \subset M$ be a smooth submanifold of M. A *distribution* D of TM on M' is a subbundle of TM over M', i.e., $D \subset pr^{-1}(M')$ with projection $pr' = pr|_D$. The annihilator of D, a codistribution on M' denoted as $D^\perp$, is a subbundle of T^*M over M' such that, for any $m \in M'$, $\omega \in D_m^\perp$ and $v \in D_m$, $\langle \omega, v \rangle = 0$. ∎

Remark 2.2.8: We make two assumptions:

(1) In the following, for simplicity, we assume $M' = M$. This assumption will be relaxed

later (after Remark 2.2.18) when we consider a special case in which M' is a leaf of a foliation of M. In this particular case, $D = TM'$ and distribution and co-distribution given below can be naturally restricted to M'.

(2) We assume that all distributions and co-distributions are smooth (i.e., differentially depend on the points of M) and nonsingular (i.e., have fixed dimension). By this assumption, it is clear that if $dim(M) = n$ and $dim(D) = d$, $dim(D^\perp) = 2n - d$. ∎

We now introduce Lagrangian systems with constraint on an n-dimensional configuration space Q. Classically, a constraint is defined by a subspace, D', of TQ, which is locally given by m independent functions, say $\phi^k(q,v)$, $k = 1,\cdots,m$, on TQ, i.e.,

$$D' = \{(q,v) \in TQ \mid \phi^k(q,v) = 0, k = 1,\cdots,m\} \qquad (2.2.16)$$

If we further assume that ϕ^k's are linear in v, D' is a distribution of TQ. Then a Lagrangian system compatible with the constraint is one such that at any point $q(t)$ on the trajectory of the motion $\{q(t), t \in [0,T], T > 0\}$, $(q(t), v(t)) = (q(t), \dot{q}(t)) \in D'$.

However, as we have seen, the geometric treatment of Lagrangian mechanics is to construct the dynamics of the system on the second tangent bundle of Q. Therefore, in general, constraint should be defined through a submanifold, or particularly, a distribution on TQ. (Caution: This is still a special case. See Remark 2.2.8 (1)). A Lagrangian system compatible with constraint will mean that the dynamics (or the special vector field) belongs to this distribution. An equivalent way to define the constraint is to consider it as a codistribution on TQ such that it annihilates the special vector fields on TQ given by the motion of the system. This is what we will do.

Definition 2.2.9: A *Constraint* on the tangent space TQ is a codistribution Θ on TQ. A Lagrangian system compatible with constraint Θ is a special vector field X on TQ, which is annihilated by Θ at every point on TQ, i.e., $\Theta(X) = 0$. ∎

In local coordinates, at each point (q,v) in TQ,

$$\Theta(q,v) \overset{\triangle}{=} span\{\theta^k(q,v) = (\alpha^k, \beta^k), \; k = 1,\cdots,m\}, \qquad (2.2.17)$$

where $\alpha^k, \beta^k \in T_q^*Q$, $k = 1,\cdots,m$, and $\{\theta^k\}$ are linearly independent, which can, therefore, be viewed as a basis for Θ.

18

Using the bundle mapping τ defined in (2.2.6) we can define the constraint reactions on TTQ and other related notions as follows.

Definition 2.2.10: The *constraint reaction* is the horizontal codistribution, $\tau\Theta$. An element of $\tau\Theta$ is referred to as a *constraint reaction force*. A constraint is *admissible* if $dim(\tau\Theta) = dim(\Theta)$. A constraint is called *ideal* if it annihilates the principal vector field X^{pr}. ∎

From (2.2.17) and (2.2.6)', in local coordinates, a constraint is admissible if $\{\beta^k, k = 1, \cdots, m\}$ are linearly independent.

The following two lemmas are essential and easy to prove [46].

Lemma 2.2.11: If a constraint is admissible, then there exists a special vector field compatible with this constraint. ∎

Lemma 2.2.12: If a constraint is ideal, constraint reaction vanishes on loops lifted on TQ from Q. ∎

Remark 2.2.13: The conditions of these two lemmas insure the existence of a special vector field which satisfies the constraint and the workless nature of constraint reaction force. Later on we will always assume these conditions are satisfied. ∎

We are now ready to state the *Lagrange-d'Alembert principle* with constraints.

Lagrange-d'Alembert Principle with Constraints 2.2.14:

On the special vector field which satisfies the constraint and determines the real trajectory of motion, the sum of the Lagrangian force and exterior force vanishes on the vector fields annihilated by the constraint reaction on the tangent bundle to the configuration space. ∎

Using the notations introduced earlier in this section, this principle says that the special vector field X which gives the dynamics of the system with constraint satisfies,

$$\langle F_L(X) + \omega_e, Z \rangle = 0 \qquad (2.2.18a)$$

for an exterior force ω_e and

$$\Theta(X) = 0 \qquad (2.2.18b)$$

for any vector field Z such that

$$(\tau\Theta)(Z) = 0. \qquad (2.2.18c)$$

Remark 2.2.15: As we have mentioned in Remark 2.2.6, vector field Z plays the role of virtual displacement. Then (2.2.18) says that work done by Lagrangian force and exterior force on virtual displacements is zero provided that the virtual displacements are "perpendicular" to the the constraint reaction. This is the essence of Lagrange-d'Alembert principle. We sometimes also call the value of Z a *test vector*. ∎

Following the same procedure used to obtain (2.2.12) and using the local expression (2.2.17) for Θ, (2.2.18) can be represented locally as

$$(\frac{d}{dt}D_2L(q,v) - D_1L(q,v)) \cdot u = \alpha_e \cdot u \qquad (2.2.19a)$$

with constraints

$$\alpha^k \cdot v + \beta^k \cdot \dot{v} = 0, \qquad k = 1, \cdots, m, \qquad (2.2.19b)$$

for u satisfying

$$\beta^k \cdot u = 0 \qquad k = 1, \cdots, m. \qquad (2.2.19c)$$

The Lagrange-d'Alembert principle can also be stated in an alternative form by using the constraint reaction force. From (2.2.18) we see that Z belongs to $(\tau\Theta)^{\perp}$ implies that $F_L(X) + \omega_e$ belongs to $\tau\Theta$, i.e., there is an element, a constraint reaction force, ω in $\tau\Theta$ such that

$$F_L(X) + \omega_e + \omega = 0. \qquad (2.2.20a)$$

Then the Lagrange-d'Alembert principle says that the special vector field X determines the real trajectory of motion for the constrained system if it satisfies (2.2.20a) and

$$\Theta(X) = 0. \qquad (2.2.20b)$$

The local expression for (2.2.20) can be obtained as follows. Let $\{\theta^k = (\alpha^k, \beta^k), k = 1, \cdots, m\}$ be the basis for Θ. Then

$$\omega = \sum_{k=1}^{m} \lambda_k(\beta^k, 0),$$

where λ_k are functions (multipliers) on TQ. Then following the same notation as for (2.2.12), the equations of motion for constrained system are

$$\frac{d}{dt}D_2L(q,v) - D_1L(q,v) = \alpha_e + \sum_{k=1}^{m} \lambda_k\beta^k \qquad (2.2.21a)$$

20

and

$$\alpha^k \cdot v + \beta^k \cdot \dot{v} = 0, \qquad k = 1, \cdots, m. \tag{2.2.21b}$$

Remark 2.2.16: In classical mechanics, both (2.2.18) and (2.2.20) (or locally, (2.2.19) and (2.2.21)) are used as dynamical equations for systems with constraints. In many physical problems, (2.2.20) is more often applied because it determines the constraint reaction force. On the other hand (2.2.18) is more important for many analytical problems. As we shell see in a later chapter where symmetry will be involved, the reduced dynamics can be derived from (2.2.18). ∎

We next consider the existence and uniqueness of a vector field on TQ which satisfies *Lagrange-d'Alembert principle*. For an unconstrained Lagrangian system, it is clear that given exterior force, there exits a unique special vector field satisfying (2.2.11) if the Lagrangian is regular, i.e., $\mathbf{F}^2 L$ is nondegenerate. Here $\mathbf{F}^2 L : TQ \to L(TQ, T^*Q)$ is the second order fiber derivative of the Lagrangian L, known as the Hessian of L. Locally, $\mathbf{F}^2 L(q, v) = D_2 D_2 L(q, v)$. For Lagrangian systems with constraint, the following proposition gives the answer. We will prove it in detail following the idea provided in [46].

Proposition 2.2.17: If the Lagrangian L is regular and the Hessian $\mathbf{F}^2 L$ is positive definite, then for every admissible constraint Θ there exists a unique special vector field X compatible with the constraint, $\Theta(X) = 0$, and satisfying the Lagrange-d'Alembert principle with constraint.

Proof: Without loss of generality, we assume there is no exterior force, i.e., $\alpha_e = 0$ (see also Remark 2.2.18(1) below). By the definition of Lagrangian force (2.2.10), we consider the existence and uniqueness of a special vector field X for equation

$$\Omega_L(X, \cdot) - dH_L + \omega = 0,$$

where Ω_L and H_L are defined respectively in (2.2.8) and (2.2.9). We prove the assertion by using local coordinates.

Since the Lagrangian L is regular, Ω_L is nondegenerate [1]. Then there exists a nondegenerate one-to-one mapping $\Pi_L : \varpi^1(TQ) \to \mathfrak{X}(TQ)$ such that

$$\Omega_L(\Pi_L(dH_L - \omega), Y) = (dH_L - \omega)(Y),$$

21

for all $Y \in \mathcal{X}(TQ)$. It is clear that $\Pi_L(dH_L - \omega)$ is special vector field since $\Pi_L(dH_L)$ is special (by Lemma 2.2.4) and $\Pi_L(\omega)$ is vertical (by Lemma 2.2.3). To bring $\Pi_L(dH_L - \omega)$ into agreement with the constraint, we require

$$\Theta(\Pi_L(dH_L - \omega)) = 0$$

or

$$\Theta(X_L) = \Theta(\Pi_L \omega),$$

where $X_L = \Pi_L(dH_L)$. If we can prove that there exists a unique $\omega \in \tau\Theta$ satisfying the above equation, the assertion is proved.

Since ω is a constraint reaction force, there exists a 1-form $\theta \in \Theta$ such that $\omega = \tau\theta$. Thus we have to solve for θ in

$$\Theta(X_L) = \Theta((\Pi_L \circ \tau)\theta). \tag{2.2.22}$$

Let $\{\theta^k\}_{k=1}^m$ be a basis for Θ. Then (2.2.22) is equivalent to

$$\langle \theta^k, X_L \rangle = \langle \theta^k, (\Pi_L \circ \tau)\theta \rangle \qquad k = 1, \cdots, m. \tag{2.2.23}$$

Let $\theta = \sum_{k=1}^m \lambda_k \theta^k$, where $\{\lambda_k\}_{k=1}^m$ are functions on TQ. Then (2.2.23) is equivalent to

$$\langle \theta^k, X_L \rangle = \sum_{i=1}^m \lambda_i \langle \theta^k, (\Pi_L \circ \tau)\theta^i \rangle, \qquad k = 1, \cdots, m. \tag{2.2.24}$$

Now we only need to show that there exists a unique solution $\{\lambda_k\}_{k=1}^m$ for the above equation, or equivalently, $m \times m$ matrix $(\langle \theta^k, (\Pi_L \circ \tau)\theta^i \rangle)$ is invertible. By the construction of operator Π_L (cf. (2.2.14)),

$$\Omega_L((\Pi_L \circ \tau)(\theta), Y) = (\tau\theta)(Y), \qquad \forall Y \in \mathcal{X}(TQ). \tag{2.2.25}$$

In local coordinates, at each point $(q, v) \in TQ$, let $\theta(q, v) = (\alpha, \beta)$ and $Y(q, v) = (Y_1, Y_2)$ where $\alpha, \beta \in T_q^* Q$ and $Y_i \in T_q Q, i = 1, 2$. We then have $\tau\theta(q, v) = (\beta, 0)$ and

$$(\Pi_L \circ \tau)(\theta)(q, v) = (0, w), \tag{2.2.26}$$

for some $w \in T_q Q$ (cf. Lemma 2.2.3). By the local expression of Ω_L (2.2.8)$'$, Equation (2.2.26) becomes

$$(D_2 D_2 L(q, v)) \cdot w \cdot Y_1 = -\beta \cdot Y_1.$$

This implies w should be, since L is regular and Y_1 is arbitrary,

$$w = -(D_2 D_2 L(q,v))^{-1}\beta.$$

Now (2.2.10) has the local expression

$$(\Pi_L \circ \tau)(\alpha,\beta) = (0, -(D_2 D_2 L)^{-1}\beta). \qquad (2.2.27)$$

Let $\theta^k(q,v) = (\alpha^k, \beta^k)$, where α^k and $\beta^k \in T_q^* Q$. Then from (2.2.27) we have

$$\langle \theta^k, (\Pi_L \circ \tau)\theta^i \rangle = -\langle \beta^k, (D_2 D_2 L(q,v))^{-1}\beta^i \rangle,$$

for $i, k = 1, \cdots, m$. Now replacing the components of right-hand-side of (2.2.24) by the right-hand-side of the above equation, we get

$$\langle \theta^k, X_L \rangle = -\sum_{i=1}^{m} \lambda_i \langle \beta^k, (D_2 D_2 L(q,v))^{-1}\beta^i \rangle, \qquad k = 1, \cdots, m. \qquad (2.2.28)$$

Note that, since Θ is admissible, the independence of $\{\theta^k\}$ implies the independence of $\{\beta^k\}$. Also, that $D_2 D_2 L(q,v)$ is positive definite implies that $(D_2 D_2 L(q,v))^{-1}$ is positive definite. Then it is easy to prove that the $m \times m$ Gramian matrix

$$(\langle \beta^k, (D_2 D_2 L(q,v))^{-1}\beta^i \rangle) \qquad (2.2.29)$$

is invertible. This completes the proof of the theorem. ■

Remark 2.2.18:

(1) In [3], the invertibility of matrix (2.2.29) is considered as the condition for the existence and uniqueness of a solution for equation (2.2.21). In classical mechanics, $L = K - V$, where K is the quadratic form of kinetic energy (or Riemannian metric on Q) and V is potential energy. Therefore the conditions of this Proposition are naturally satisfied.

(2) If the exterior force or control $\omega_e \neq 0$, (2.2.24) can be written as

$$\langle \theta^k, \Pi_L(dH_L - \omega_e) \rangle = \sum_{i=1}^{m} \lambda_i \langle \theta^k, (\Pi_L \circ \tau)\theta^i \rangle, \qquad k = 1, \cdots, m,$$

and the solvability conditions is as before. To keep the system moving under the constraints, one should make $dH_L - \omega_e \neq 0$. ■

We now study a special type of constraint, which we briefly introduced earlier after Remark 2.2.8. Consider a foliation of TQ given, locally, by a set of smooth and independent functions

$$\{\phi^k : TQ \to \Re, k = 1, \cdots, m\}. \tag{2.2.30}$$

We would like to study the motion of a Lagrangian system on a leaf, say zero level leaf, of this foliation,

$$\mathcal{S} = \{(q, v) \in TQ | \phi^k(q, v) = 0, k = 1, \cdots, m\}. \tag{2.2.31}$$

Then the constraint Θ is, locally,

$$\Theta(q, v) = span\{\theta^k = (D_1\phi^k(q, v), D_2\phi^k(q, v)), k = 1, \cdots, m\}. \tag{2.2.32}$$

Remark 2.2.19: One should note that this constraint, or co-distribution, is defined on $\mathcal{S}$, instead of all of TQ. As we have mentioned in Remark 2.2.8(1), the geometric objects (distributions and vector fields) in the above Definitions, Lagrange-d'Alembert principle and the Proposition can be modified (restricted) to $\mathcal{S}$ with impugnity. ∎

By Lagrange-d'Alembert principle and (2.2.32), the restricted dynamic motion on $\mathcal{S}$, satisfies the following equations

$$(\frac{d}{dt}D_2L(q, v) - D_1L(q, v)) \cdot u = \alpha_e \cdot u \tag{2.2.33a}$$

for u satisfying

$$D_2\phi^k(q, v) \cdot u = 0, \qquad k = 1, \cdots, m \tag{2.2.33b}$$

and (q, v) satisfying

$$\phi^k(q, v) = 0, \qquad k = 1, \cdots, m; \tag{2.2.33c}$$

or, using multipliers,

$$\frac{d}{dt}D_2L(q, v) - D_1L(q, v) = \alpha_e + \sum_{k=1}^{m} \lambda_k D_2\phi^k(q, v) \tag{2.2.34a}$$

and

$$\phi^k(q, v) = 0, \qquad k = 1, \cdots, m. \tag{2.2.34b}$$

A further special case of the above constraint is that the ϕ^k's are linear functions of v, i.e.,

$$\phi^k(q, v) = \langle \omega^k(q), v \rangle$$

for some $\omega^k \in \varpi^1(Q')$ for $k = 1, \cdots, m$ and a submanifold Q' of Q. With this construction, the constraints can also be considered as a co-distribution on Q', denoted as

$$\mathcal{B} = \{\omega^k, k = 1, \cdots, m\}.$$

In this case, the equations of motion are (2.2.33) or (2.2.34) with $D_2\phi^k = \omega^k, k = 1, \cdots, m$.

Definition 2.2.20: If the co-distribution $\mathcal{B}$ is not integrable, it is called a *nonholonomic constraint* on Q', otherwise it is a *holonomic constraint* on Q'. ∎

Remark 2.2.21: Equivalently, one can use the distribution $\mathcal{B}^\perp$ to define nonholonomic and holonomic constraints on Q'. ∎

Remark 2.2.22: In holonomic problems, constraints are usually given by a set of functions on Q, say

$$\{f^k : Q \to \Re, k = 1, \cdots, m\}, \tag{2.2.35}$$

which gives a foliation of the configuration space Q. Then, the submanifold Q' in Definition 2.2.20 is a leaf of this foliation. From Lagrange-d'Alembert principle, the motion of the system satisfies the following equations:

$$\frac{d}{dt}D_2L(q, v) - D_1L(q, v) = \alpha_e + \sum_{k=1}^{m} \lambda_k Df^k(q) \tag{2.2.36a}$$

and

$$f^k(q) = 0, \qquad k = 1, \cdots, m. \tag{2.2.36b}$$

From a mathematical point of view, the holonomic problem is viewed as a simple case since it can be re-arranged as an unconstrained problem (at least locally). But, as we will see in next chapter, physical systems (such as four-bar linkages) with holonomic constraints may possess very interesting properties in their kinematics and dynamics. ∎

Remark 2.2.23: It is known that for the systems with holonomic constraints, the *Lagrange-d'Alembert's principle* is equivalent to the (Hamilton's) principle of least action, which determines a curve, $\{q(t), t \in [0, 1]\}$, in Q to satisfying two fixed end points and constraints, and minimizes the functional

$$\int_0^1 L(q(t), \dot{q}(t))dt,$$

where L is the Lagrangian. It is also known that, in general, if the constraints are nonholonomic, the two principles are not equivalent. In particular, if the leaf constraints are given by nonintegrable functions (2.2.30), the curve $q(\cdot) \in Q$ determined by the *principle of least action with constraints* fulfills the equation

$$\frac{d}{dt} D_2 L(q,v) - D_1 L(q,v) = \sum_{k=1}^{m} \lambda_k (\frac{d}{dt} D_2 \phi^k(q,v) - D_1 \phi^k(q,v)) + \sum_{k=1}^{m} \dot{\lambda}_k D_2 \phi^k(q,v)$$

(2.2.37a)

or, equivalently,

$$\frac{d}{dt} D_2 \mathcal{L}(q,v) - D_1 \mathcal{L}(q,v) = 0,$$

(2.2.37b)

where $\mathcal{L}(q,v) = L(q,v) - \sum_{k=1}^{m} \lambda_k \phi^k(q,v)$. In [3] the systems determined by the equations (2.2.37) are said to obey *vakonomic mechanics* (mechanics of variational axiomatic kind). It has been shown that for mechanical systems with nonholonomic constraints, only the Lagrange-d'Alembert principle provides physical motion (cf. [39, 40,48]). If Q is Riemannian manifold with L the metric, the above assertions say that for holonomic mechanics, the real motions of the system are geodesics, but for nonholonomic mechanics, it is not. In Chapter IV we will study a special nonholonomic variational problem with a meaningful choice of Riemannian metric. ∎

Example 2.2.24: Before ending this chapter, we consider a popular example in the literature of nonholonomic dynamics, that is, *rolling (without sliding) a homogeneous sphere on an absolutely rough and horizontal plane*. We will solve the constrained dynamic equation involving test vectors.

Let m be the mass of the sphere and mk^2 the inertia of the sphere about any axis passing through the center, where k is called radius of gyration. Let a be the radius of the sphere. Let $0\text{-}XYZ$ be a coordinate system in inertial space with Z-axis being perpendicular to the plane. Since it is obvious that there is no motion along Z direction, the configuration space of the system is $\Re^2 \times SO(3)$. The Lagrangian of the system can be expressed as

$$L = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2} mk^2(\omega_x^2 + \omega_y^2 + \omega_z^2),$$

(2.2.38)

where (x,y) gives the location of the center of the sphere on the plane and $(\omega_x, \omega_y, \omega_z)$ is the angular velocity of the sphere expressed in coordinate system $0\text{-}XYZ$. The

nonholonomic constraint of rolling without sliding can be expressed as

$$a\omega_x = \dot{y} \quad \text{and} \quad a\omega_y = -\dot{x}. \tag{2.2.39}$$

We also assume that the sphere rotates on the plane under the applied force (F_x, F_y) acting at the center of the sphere. With the above setting, using (2.2.33), we get the following dynamic equations with constraints:

$$m(\ddot{x}u_1 + \ddot{y}u_2) + mk^2(\dot{\omega}_x\mu_1 + \dot{\omega}_y\mu_2 + \dot{\omega}_z\mu_3) = F_x u_1 + F_y u_2 \tag{2.2.40a}$$

for $(u_1, u_2, \mu_1, \mu_2, \mu_3)$ being test vector satisfying

$$a\mu_1 = u_2 \quad \text{and} \quad a\mu_2 = -u_1 \tag{2.2.40b}$$

and $(\dot{x}, \dot{y}, \omega_x, \omega_y, \omega_z)$ satisfying linear constraints

$$a\omega_x = \dot{y} \quad \text{and} \quad a\omega_y = -\dot{x}. \tag{2.2.40c}$$

After a simple calculation to eliminate μ_1 and μ_2, one can get

$$m\frac{a^2 + k^2}{a^2}(\ddot{x}u_1 + \ddot{y}u_2) + mk^2\dot{\omega}_z\mu_3 = F_x u_1 + F_y u_2, \tag{2.2.41}$$

where u_1, u_2 and μ_3 can be chosen arbitrarily. Then from (2.2.41) and constraints (2.2.40c), one gets

$$\begin{cases} \ddot{x} = \dfrac{a^2}{a^2 + k^2}\dfrac{F_x}{m} \\[2mm] \ddot{y} = \dfrac{a^2}{a^2 + k^2}\dfrac{F_y}{m} \\[2mm] \dot{\omega}_x = \dfrac{a}{a^2 + k^2}\dfrac{F_y}{m} \\[2mm] \dot{\omega}_y = -\dfrac{a}{a^2 + k^2}\dfrac{F_x}{m} \\[2mm] \dot{\omega}_z = 0. \end{cases} \tag{2.2.42}$$

If F_x, F_y is known the above equation can be easily solved.

There are two important features in the above equations. First, the motion of the center of the sphere does not depend on the spin of the sphere about vertical direction of the plane, i.e., the motion of the center looks like that of a particle with modified mass. Secondly, angular velocity ω_z is a constant. An interesting question one may ask is that

27

if we consider conserved ω_z as an *a priori* constraint, by adding it to the constraints in (2.2.39) and using the above approach, can we get the same dynamic equation? The answer is yes. In fact, equations (2.2.40) now have the following form:

$$m(\ddot{x}u_1 + \ddot{y}u_2) + mk^2(\dot{\omega}_x\mu_1 + \dot{\omega}_y\mu_2 + \dot{\omega}_z\mu_3) = F_x u_1 + F_y u_2 \qquad (2.2.43\text{a})$$

for $(u_1, u_2, \mu_1, \mu_2, \mu_3)$ being test vector satisfying

$$a\mu_1 = u_2 \quad \text{and} \quad a\mu_2 = -u_1 \quad \text{and} \quad \mu_3 = 0 \qquad (2.2.43\text{b})$$

and $(\dot{x}, \dot{y}, \omega_x, \omega_y, \omega_z)$ satisfying constraints

$$a\omega_x = \dot{y} \quad \text{and} \quad a\omega_y = -\dot{x} \quad \text{and} \quad \omega_z = c, \qquad (2.2.43\text{c})$$

which is affine. After a straightforward calculation from (2.2.43), one gets

$$m\frac{a^2 + k^2}{a^2}(\ddot{x}u_1 + \ddot{y}u_2) = F_x u_1 + F_y u_2, \qquad (2.2.44)$$

where u_1, u_2 can be chosen arbitrarily. Using (2.2.44) and (2.2.43c), the final result is the same (2.2.42).

It is very important to observe that, after using $\omega_z = c$ as one of the constraints, we get a lower dimensional dynamic equation (2.2.44) which is independent of rotational variables. The theory behind this phenomenon and the systematic approach to derive such reduced equations are precisely what we will explore in Chapter V. ∎

CHAPTER III

GEOMETRY AND DYNAMICS OF FLOATING FOUR-BAR LINKAGES

Inspired by many mechanical problems in space engineering, significant progress has been made to understand dynamics of coupled mechanical systems, which consist of *serial links* or *open chains*, by applying geometric methods, symmetry principles and reduction, cf. [21] and references therein. On the other hand, problems in aerospace engineering also suggest that multibody systems with kinematic loops are of practical importance. In this chapter, we study the kinematics and dynamics of the simplest mechanical system with a kinematic loop, which is represented as (kinematic) holonomic constraint — the planar, *floating* four-bar linkage. By *floating* we mean that no link is fixed in space. It is simplest in that it has the fewest degrees of freedom among all kinds of mechanisms with closed loops.

There are at least two important aspects of our study of such systems. First, as we shall see in this chapter, the presence of loop closure constraints implies that the knowledge of the Hamiltonian structure and phase portraits for open chain multibody systems cannot be applied directly to systems with closed loops. A careful study of a simple coupled system with closed loop is necessary. Secondly, we note that in the

theory of machines, four-bar linkage (with one link, the *frame*, fixed) are often used to synthesize more complex mechanisms [19]. Moreover, it can generate the wide variety of motions represented by *coupler curves* [18]. For *floating* four-bar linkages, one expects similar properties to hold.

The outline of this chapter is as follows. After stating the basic notations for this chapter in Section 3.1, we give a geometric description of the configuration space in Section 3.2. In Section 3.3, we give an explicit expression for the kinetic energy of the system. In Section 3.4, we explore symmetry properties, Hamiltonian structure and reduction of four-bar linkage dynamics. In Section 3.5 a theorem of Smale is used to compute relative equilibria for the dynamics of a four-bar linkage. Then, using techniques from singularity theory, we study the local bifurcations of relative equilibria for linkages which admit symmetric shapes.

3.1 Notations and Geometric Constraints

The structure of a closed floating four-bar linkage is represented in Figure 3.1.1. What we mean by *bar* in this chapter is a planar rigid body, on which the center of mass and pin joints are located arbitrarily. The bars are labeled sequentially from 0 to 3. On each bar, a body-fixed frame is attached such that its origin is at the center of mass of the bar and the x-axis is parallel to the line connecting two joints on the bar. In particular, we choose the positive direction of of x-axis of i-th bar towards the $(i+1)$-th bar, for $i = 0, 1, 2, 3 \pmod 4$. We define the following quantities.

$\mathbf{d}_{ij}$ the vector from the center of mass of i-th bar to the joint with j-th bar in body-fixed frame;

$\mathbf{r}_i$ the position vector of the center of mass of i-th bar relative to an inertial observer;

$\mathbf{r}_i^c$ the vector from the system center of mass to the center of mass of i-th bar;

$\mathbf{r}_c$ the position of the system center of mass relative to the reference point of the inertial observer;

<div align="center">30</div>
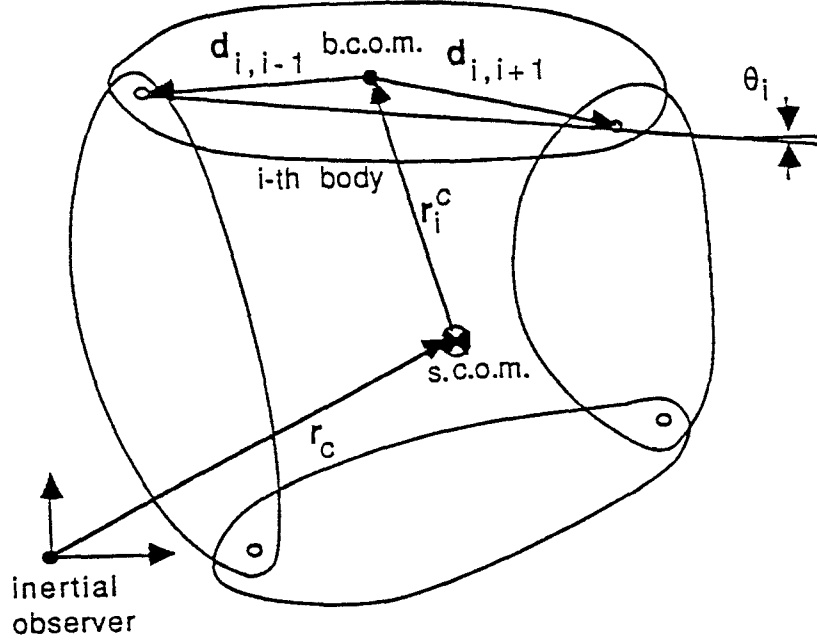
Figure 3.1.1 The general structure of four-bar linkage

θ_i the orientation angle of i-th bar relative to the inertial frame;

θ_{ij} the relative angle between i-th bar and j-th bar, i.e., $\theta_{ij} = \theta_i - \theta_j$;

l_i the *length* of i-th bar, which is defined as the distance between the joints
on i-th bar, i.e., $l_i = \|d_{i,i+1} - d_{i,i-1}\|$;

m_i, I_i the mass and the moment of inertia of i-th bar about its center of mass;

m the total mass of the system, i.e.,

$$m = \sum_{i=0}^{3} m_i.$$

Any pair of adjacent bodies is connected by the following relation, *the hinge constraint*,

$$\mathbf{r}_{i+1}^c = \mathbf{r}_i^c + R(\theta_i)\mathbf{d}_{i,i+1} - R(\theta_{i+1})\mathbf{d}_{i+1,i} \qquad i = 0,1,2,3 \pmod 4, \qquad (3.1.1)$$

31

where

$$R(\theta_i) = \begin{pmatrix} cos(\theta_i) & -sin(\theta_i) \\ sin(\theta_i) & cos(\theta_i) \end{pmatrix}$$

the rotation matrix. By eliminating $\mathbf{r}_i^c$ in (3.1.1) we find the loop constraint or *closure constraint*,

$$\sum_{i=0}^{3} R(\theta_i)(\mathbf{d}_{i,i+1} - \mathbf{d}_{i,i-1}) = 0, \tag{3.1.2}$$

where the convention of modulo 4 addition for the subscripts is adopted.

3.2 The Configuration Space

In this section we investigate the conditions under which the loop constraint (3.1.2) describes a submanifold of the configuration manifold of an open four-bar chain.

For a planar floating four-bar *open* chain, the configuration space is

$$M = \Re^2 \times S^1 \times S^1 \times S^1 \times S^1.$$

Here M is a 6-dimensional manifold with local coordinates $\mathbf{m} = (x_0, y_0, \theta_0, \theta_1, \theta_2, \theta_3)$. This corresponds to keeping track of a material point (say center of mass on one of the bodies) and the four absolute orientations. See [34,43,44] for the Hamiltonian mechanics of such open chains.

For a *closed* four-bar mechanism considered in this chapter, the configuration space denoted by Q is a subset of M determined by (3.1.2), i.e.,

$$Q = \{\mathbf{m} \in M | \mathbf{F}(\mathbf{m}) = 0\}$$

for

$$\mathbf{F}(\mathbf{m}) = \sum_{i=0}^{3} R(\theta_i)(\mathbf{d}_{i,i+1} - \mathbf{d}_{i,i-1}). \tag{3.2.1}$$

For Q to be a submanifold of M, we have the following condition.

Theorem 3.2.1: If

$$l_0 \pm l_1 \pm l_2 \pm l_3 \neq 0,$$

Q is a submanifold of M.

Proof: Note that $\mathbf{F} : M \longrightarrow \Re^2$. From [17] we know that if $\mathbf{0}$ is a regular value of $\mathbf{F}$, i.e., $\partial \mathbf{F}/\partial \mathbf{m}$ has full rank for all $\mathbf{m} \in M$ satisfying $\mathbf{F}(\mathbf{m}) = \mathbf{0}$, then Q is a submanifold of M.

From (3.2.1) we have

$$\frac{\partial \mathbf{F}}{\partial \mathbf{m}} = \begin{pmatrix} 0 & 0 & -l_0 sin(\theta_0) & -l_1 sin(\theta_1) & -l_2 sin(\theta_2) & -l_3 sin(\theta_3) \\ 0 & 0 & l_0 cos(\theta_0) & l_1 cos(\theta_1) & l_2 cos(\theta_2) & l_3 cos(\theta_3) \end{pmatrix}.$$

It is easy to check that all the nontrivial determinants of 2×2 submatrices are given by the following functions,

$$g_1(\mathbf{m}) = l_0 l_1 sin(\theta_1 - \theta_0)$$

$$g_2(\mathbf{m}) = l_0 l_2 sin(\theta_2 - \theta_0)$$

$$g_3(\mathbf{m}) = l_0 l_3 sin(\theta_3 - \theta_0) \qquad (3.2.2)$$

$$g_4(\mathbf{m}) = l_1 l_2 sin(\theta_2 - \theta_1)$$

$$g_5(\mathbf{m}) = l_1 l_3 sin(\theta_3 - \theta_1)$$

$$g_6(\mathbf{m}) = l_2 l_3 sin(\theta_3 - \theta_2).$$

Therefore, if for each $\mathbf{m} \in M$ satisfying $\mathbf{F}(\mathbf{m}) = 0$, there exists an i such that $g_i(\mathbf{m}) \neq 0$, Q is a submanifold of M. It is obvious that the above condition depends on the relative angles and, consequently, the lengths of the bars. To arrive at the condition in the statement of the theorem, we proceed from the converse.

If $g_i(\mathbf{m}) = 0$ *for all* i, from (3.2.2), we have

$$\theta_1 - \theta_0 = 0 \; or \; \pi \qquad (3.2.3a)$$

and

$$\theta_3 - \theta_2 = 0 \; or \; \pi \qquad (3.2.3b)$$

and

$$\theta_3 - \theta_0 = 0 \; or \; \pi. \qquad (3.2.3c)$$

33

Premultiplying (3.1.2) by $R(-\theta_0)$, we get following equivalent closure constraint equations

$$l_0 + l_1 cos(\theta_1 - \theta_0) + l_2(cos(\theta_3 - \theta_2)cos(\theta_3 - \theta_0)$$

$$+ sin(\theta_3 - \theta_2)sin(\theta_3 - \theta_0)) + l_3 cos(\theta_3 - \theta_2) = 0 \qquad (3.2.4a)$$

$$l_1 sin(\theta_1 - \theta_0) + l_2(cos(\theta_3 - \theta_2)sin(\theta_3 - \theta_0)$$

$$- sin(\theta_3 - \theta_2)cos(\theta_3 - \theta_0)) + l_3 sin(\theta_3 - \theta_2) = 0. \qquad (3.2.4b)$$

So, to make θ_i's satisfy both (3.2.3) and (3.2.4), the lengths of the bars should satisfy

$$l_0 + (-1)^{k_1} l_1 + (-1)^{k_2} l_2 + (-1)^{k_3} l_3 = 0 \qquad (3.2.5)$$

for some $k_1, k_2, k_3 \in \{0,1\}$. By contradiction, if, *for all* $k_1, k_2, k_3 \in \{0,1\}$, (3.2.5) does not hold, then there is an i such that $g_i(\mathbf{m}) \neq \mathbf{0}$, which means that Q is a submanifold of M. ∎

Table 3.2.1 summarizes the cases which violate the conditions on lengths of the bars in Theorem 3.2.1. It is easy to observe that case (i) can never happen since l_i are assumed to be positive. In addition, cases (iii), (iv), (v) and (vi) are trivial since, in these cases, none of the relative angles can vary, i.e., the configuration space loses one degree of freedom and the linkage becomes a rigid structure.

In the following we show that the condition in Theorem 3.2.1 can be simplified by ignoring the labels on the bars and give a topological description of Q. We first recall some definitions and results in the classical theory of mechanisms [15,36,37].

For a four-bar linkage in classical mechanics, in which one bar, *the frame*, is fixed, the following quantities are defined:

$$\bar{s} = \text{length of shortest bar,}$$

$$\bar{l} = \text{length of longest bar,}$$

$$\bar{p}, \bar{q} = \text{lengths of intermediate bars.}$$

A bar which is free to rotate through 2π with respect to a second bar is said to *revolve* relative to the second bar and is referred to as a *crank*. Any bar which does not revolve
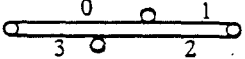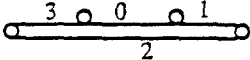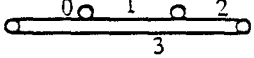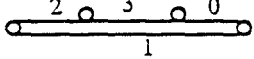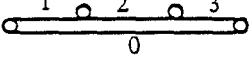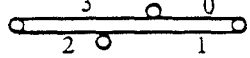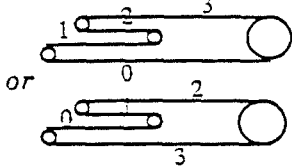
case	$\theta_1-\theta_0$	$\theta_3-\theta_2$	$\theta_3-\theta_0$	link condition	structure
i	0	0	0	$.l_0+l_1+l_2+l_3=0$	
ii	0	0	π	$l_0+l_1-l_2-l_3=0$	
iii	0	π	0	$l_0+l_1-l_2+l_3=0$	
iv	0	π	π	$l_0+l_1+l_2-l_3=0$	
v	π	0	0	$l_0-l_1+l_2+l_3=0$	
vi	π	0	π	$l_0-l_1-l_2-l_3=0$	
vii	π	π	0	$l_0-l_1-l_2+l_3=0$	
viii	π	π	π	$l_0-l_1+l_2-l_3=0$	

Table 3.2.1 Singular Configurations

is called a *rocker*. If it is possible for all bars to become simultaneously aligned, such a state is called a *change point* and the linkage is said to be a *change-point mechanism*. Given four bars, assembling them into a closed loop and labeling them sequentially, the linkage may have instantaneously one of three kinds of shape. These correspond to

$$sin(\theta_3 - \theta_2) \begin{cases} > 0; \\ < 0; \\ = 0. \end{cases}$$

In the classical theory of mechanisms, they are called *lagging form, leading form* and *dead point*, respectively. The following theorem is due to Grashof [15].

Theorem 3.2.2:

(1) A four-bar mechanism has at least one crank if

$$\bar{s} + \bar{l} \leq \bar{p} + \bar{q}$$

and all three will rock if

$$\bar{s} + \bar{l} > \bar{p} + \bar{q}.$$

(2) A four-bar mechanism is a change-point mechanism if and only if

$$\bar{s} + \bar{l} = \bar{p} + \bar{q}. \qquad \blacksquare$$

Remarks 3.2.3:

(1) If $\bar{s} + \bar{l} < \bar{p} + \bar{q}$, the shortest bar is the revolving bar.

(2) It is easy to check that the cases (ii), (vii) and (viii) in Table 3.2.1 correspond to

$\bar{s} + \bar{l} = \bar{p} + \bar{q}$, i.e., they correspond to change-point mechanisms. $\qquad \blacksquare$

Since the conditions in Grashof's Theorem 3.2.2 are so important, we give the following definition.

Definition 3.2.4: We refer to the condition $\bar{s} + \bar{l} < \bar{p} + \bar{q}$ and $\bar{s} + \bar{l} > \bar{p} + \bar{q}$ as *Grashof* and *non-Grashof condition*, respectively. The corresponding linkages are referred to as *Grashof* and *non-Grashof mechanism* or *linkage*, respectively. $\qquad \blacksquare$

From Theorem 3.2.1, we immediately have following result.

Corollary 3.2.5: If a four-bar linkage is constructible, i.e., $\bar{l} \le \bar{s} + \bar{q} + \bar{p}$ and

$$\bar{s} + \bar{l} \ne \bar{p} + \bar{q},$$

Q is a submanifold of M. In other words, Q is a submanifold of M if either Grashof or non-Grashof condition holds. $\qquad \blacksquare$

To give a topological description of the configuration manifold Q, we first need the following result.

Proposition 3.2.6: Let $l_1 = \bar{s}$. Then $\bar{s} + \bar{l} < \bar{p} + \bar{q}$ if and only if $sin(\theta_3 - \theta_2) \ne 0$ for all configurations.

Proof: The mechanism can be assembled with $\bar{s}$ adjacent to $\bar{l}$ or with $\bar{s}$ opposite $\bar{l}$. And, $\bar{l}$ can be l_0, l_2 or l_3. If $\theta_3 - \theta_2 = k\pi$, the whole structure attains a triangular shape which has the property that the sum of two sides is larger than the third one. Then it is easy to check that all possible cases will lead to $\bar{s} + \bar{l} > \bar{p} + \bar{q}$. Further, it is obvious that if $\bar{s} + \bar{l} = \bar{p} + \bar{q}$, there exist θ_2 and θ_3 such that $sin(\theta_3 - \theta_2) = 0$. And, if $\bar{s} + \bar{l} > \bar{p} + \bar{q}$, by Grashof's theorem, all bars will rock with respect to each other. This

means that dead point is reachable. Therefore, if $sin(\theta_3 - \theta_2) \neq 0$ for all configurations, $\bar{s} + \bar{l} < \bar{p} + \bar{q}$. ∎

An equivalent way to state the above assertion is that a four-bar linkage is a Grashof mechanism if and only if it is constrained to be in either leading form or lagging form. Moreover, for non-Grashof mechanisms, the linkage can vary continuously from leading form to lagging form.

From Grashof's theorem and the above proposition we can get a topological description for Q.

Theorem 3.2.7:

(a) For a Grashof linkage, i.e., $\bar{s} + \bar{l} < \bar{p} + \bar{q}$, Q has two components. Each component is diffeomorphic to $\Re^2 \times S^1 \times S^1$.

(b) For a non-Grashof linkage, i.e., $\bar{s} + \bar{l} > \bar{p} + \bar{q}$, $Q = \Re^2 \times S^1 \times S^1$.

Proof: Our proof is based on explicit parameterization of the manifold Q. Recall that from (3.2.1) the dimension of Q is four. Again, we let $\bar{s} = l_1$.

(a) If $\bar{s} + \bar{l} < \bar{p} + \bar{q}$, we consider the parameters $(x_0, y_0, \theta_0, \theta_1)$, where (x_0, y_0) is the coordinate of any point on 0-th bar in inertial frame. From the definitions of θ_0 and θ_1 and the Grashof Theorem 3.2.2, both θ_0 and θ_1 can vary from $-\pi$ to π independently. From Proposition 3.2.8, one component of the manifold Q corresponds to leading form. The other corresponds to lagging form.

(b) If $\bar{s} + \bar{l} > \bar{p} + \bar{q}$, from Grashof theorem, there exists an angle α, $0 < \alpha < \pi$, such that, $\theta_{10} \in [-\alpha + \pi, \alpha + \pi]$. At the boundaries, the system is at "dead point". Now, we consider the independent parameters $(x_0, y_0, \theta_0, \hat{\theta})$, where

$$\hat{\theta} = \begin{cases} -\frac{\theta_{10} - \pi - \alpha}{\alpha} \frac{\pi}{2}, & \text{if } sin(\theta_{32}) \geq 0 \text{ (lagging form)}; \\ \frac{\theta_{10} - \pi - \alpha}{\alpha} \frac{\pi}{2}, & \text{if } sin(\theta_{32}) < 0 \text{ (leading form)}. \end{cases} \tag{3.2.6}$$

It is easy to check that $\hat{\theta}$ and θ_{10} have a one-to-one relation in both leading form and lagging form. Moreover, $\hat{\theta}$ varies from $-\pi$ to π. For $\hat{\theta} \in (-\pi, 0)$, the mechanism is in the leading form; for $\hat{\theta} \in [0, \pi]$, the mechanism is in lagging form. So $(x_0, y_0, \theta_0, \hat{\theta})$ parameterize $\Re^2 \times S^1 \times S^1$. ∎

Remarks 3.2.8:

(1) Gibson and Newstead in [13] proved above assertion using the methods of algebraic geometry.

(2) This theorem says that for a fixed location of the center of mass in $\Re^2$, the configuration space is the disjoint union of two tori for a Grashof linkage and a torus for a non-Grashof linkage. For the latter, the torus can be split into two parts, one corresponding to leading form, the other corresponding to lagging form. It can be imagined that, when $\bar{s} + \bar{l} = \bar{q} + \bar{p}$, two tori touch each other with a circular intersection. This can be seen in the following example.

(3) From the above remark, we note that, for both Grashof and non-Grashof linkages, one *may* use $(x_0, y_0, \theta_0, \theta_1)$ and specified form (leading or lagging) to parameterize Q *locally*. For the non-Grashof case, one has to worry about the parameterization in the neighborhood of the dead point. This problem can be solved by re-labeling the bars. We will discuss it further in the next section. ∎
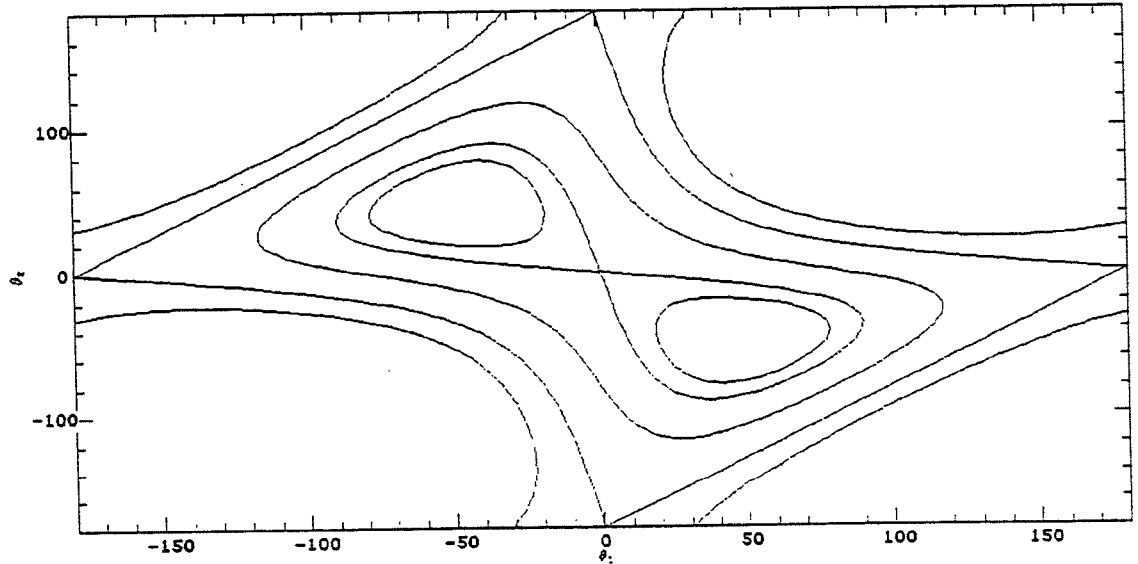
Figure 3.2.1 Reduced configuration spaces, an example

Example 3.2.9: Here we illustrate the Theorems 3.2.1 and 3.2.7 by a simple example. First, we fix one bar (position and orientation), say l_0 and let $\theta_0 = 0$. This means a point in $\Re^2 \times S^1$ has been chosen and the dimension of the configuration is reduced to one. Now the constraint equation (3.1.2) becomes

$$l_0 + l_1 cos(\theta_1) + l_2 cos(\theta_2) + l_3 cos(\theta_3) = 0$$

$$l_1 sin(\theta_1) + l_2 sin(\theta_2) + l_3 sin(\theta_3) = 0.$$

Eliminating one more angle, say θ_3, we get

$$f(\theta_1, \theta_2) \triangleq (l_0 + l_1 cos(\theta_1) + l_2 cos(\theta_2))^2 + (l_1 sin(\theta_1) + l_2 sin(\theta_2))^2 - l_3^2 = 0. \quad (3.2.7)$$

The solutions of this equation gives a curve on a torus T^2 as the configuration space.

Now we choose $l_0 = 3$, $l_1 = 3$, $l_3 = 4$ and let l_2 vary. Figure 3.2.1 shows the results. The rectangle with opposite edges identified is the standard way to represent a torus [17]. We see that when $l_2 = 1.9$, $\bar{s} + \bar{l} < \bar{p} + \bar{q}$ holds and the solution of (3.2.7) on T^2 is a disconnected closed curve; when $l_2 = 2$, $\bar{s} + \bar{l} = \bar{p} + \bar{q}$ holds and the solution of (3.2.7) on T^2 is a "figure 8"; when $l_2 = 3$, $\bar{s} + \bar{l} > \bar{p} + \bar{q}$ holds and the solution of (3.2.7) on T^2 is a connected closed curve; when $l_2 = 4$, $\bar{s} + \bar{l} = \bar{p} + \bar{q}$ holds and the solution of (3.2.7) on T^2 is a "figure 8", again; when $l_2 = 4.5$, $\bar{s} + \bar{l} > \bar{p} + \bar{q}$ holds and the solution of (3.2.7) on T^2 is a single closed curve, again. If the linkage is allowed to float, we get the configuration spaces described in Theorem 3.2.7. ∎

3.3 Kinetic Energy

In this section we derive the kinetic energy, or Lagrangian since we assumed that no potential energy is involved, for the whole system. The basic idea is to write the kinetic energy for each individual body first and then use the constraint equations to eliminate extra variables.

The kinetic energy of the i-th bar is

$$T_i = \frac{1}{2} \omega_i^2 I_i + \frac{1}{2} m_i \|\dot{\mathbf{r}}_i\|^2,$$

39

where $\omega_i = \dot{\theta}_i$. The total kinetic energy is

$$T = \frac{1}{2} \sum_{i=0}^{3} \omega_i^2 I_i + \frac{1}{2} \sum_{i=0}^{3} m_i \|\dot{\mathbf{r}}_i\|^2. \tag{3.3.1}$$

To describe the kinetic energy relative to the center of mass, we have following useful equations,

$$\mathbf{r}_i = \mathbf{r}_c + \mathbf{r}_i^c, \qquad i = 0, 1, 2, 3 \tag{3.3.2}$$

$$\sum_{i=0}^{3} m_i \mathbf{r}_i^c = 0. \tag{3.3.3}$$

By applying (3.3.2) and (3.3.3), (3.3.1) becomes,

$$T = \frac{1}{2} \sum_{i=0}^{3} \omega_i^2 I_i + \frac{1}{2} \sum_{i=0}^{3} m_i \|\dot{\mathbf{r}}_i^c\|^2 + \frac{1}{2} m \|\dot{\mathbf{r}}_c\|^2. \tag{3.3.4}$$

Applying (3.1.1) and (3.3.2), we get

$$
\begin{aligned}
\mathbf{r}_i^c = \frac{1}{m} \big[& R(\theta_{i-1}) m_{i-1} \mathbf{d}_{i-1,i} \\
& - R(\theta_i)(m_{i-1}\mathbf{d}_{i,i-1} + (m_{i+1} + m_{i+2})\mathbf{d}_{i,i+1}) \\
& + R(\theta_{i+1})((m_{i+1} + m_{i+2})\mathbf{d}_{i+1,i} - m_{i+2}\mathbf{d}_{i+1,i+2}) \\
& + R(\theta_{i+2}) m_{i+2} \mathbf{d}_{i+2,i+1} \big],
\end{aligned}
$$

for $i = 0, 1, 2, 3 \pmod 4$. Note that the convention on the subscript can be used because of the closed-loop condition. Furthermore,

$$
\begin{aligned}
\dot{\mathbf{r}}_i^c = \frac{1}{m} \big[& m_{i-1}\hat{\omega}_{i-1} R(\theta_{i-1})\mathbf{d}_{i-1,i} \\
& - \hat{\omega}_i R(\theta_i)(m_{i-1}\mathbf{d}_{i,i-1} + (m_{i+1} + m_{i+2})\mathbf{d}_{i,i+1}) \\
& + \hat{\omega}_{i+1} R(\theta_{i+1})((m_{i+1} + m_{i+2})\mathbf{d}_{i+1,i} - m_{i+2}\mathbf{d}_{i+1,i+2}) \\
& + \hat{\omega}_{i+2} R(\theta_{i+2}) m_{i+2} \mathbf{d}_{i+2,i+1} \big],
\end{aligned}
\tag{3.3.5}
$$

for $i = 0, 1, 2, 3 \pmod 4$, where

$$\hat{\omega}_i = \begin{pmatrix} 0 & -\omega_i \\ \omega_i & 0 \end{pmatrix}.$$

By substituting the formula for $\dot{\mathbf{r}}_i^c$ into (3.3.4), we get a more compact expression for the kinetic energy

$$T = \frac{1}{2} < \tilde{\omega}, \tilde{M}\tilde{\omega} > +\frac{1}{2}m\|\dot{r}_c\|^2, \tag{3.3.6}$$

where $\tilde{\omega} = (\omega_0, \omega_1, \omega_2, \omega_3)^T$ and $\tilde{M} = (\tilde{M}_{i,j}, i, j = 0, 1, 2, 3)$ is a 4×4 symmetric matrix, with elements given as follows.

Let

$$\bar{m}_i^I = \frac{1}{m^2}[m_i m_{i+1}(m_i + m_{i+1})$$

$$+ m_{i+1}(m_{i+1}m_{i-1} + m_i m_{i+2})$$

$$+ m_i m_{i+2}(m_{i+1} + m_{i+2})] \tag{3.3.7a}$$

$$\bar{m}_i^{II} = \frac{m_i}{m^2}(m_{i+2}^2 - m_{i+1}m_{i-1}) \tag{3.3.7b}$$

$$\bar{m}_i^{III} = \frac{m_i m_{i+2}}{m^2}(m_{i+1} + m_{i-1}). \tag{3.3.7c}$$

Then

$$\tilde{M}_{ii} = I_i + \bar{m}_i^I \|d\|_{i,i+1}^2 + \bar{m}_{i-1}^I \|d\|_{i,i-1}^2$$

$$- 2\bar{m}_{i-1}^{II} < d_{i,i+1}, d_{i,i-1} > \tag{3.3.8a}$$

$$\tilde{M}_{i,i+1} = -\bar{m}_i^I < d_{i,i+1}, R(\theta_{i+1,i})d_{i+1,i} >$$

$$+ \bar{m}_i^{II} < d_{i,i+1}, R(\theta_{i+1,i})d_{i+1,i+2} >$$

$$+ \bar{m}_{i-1}^{II} < d_{i,i-1}, R(\theta_{i+1,i})d_{i+1,i} >$$

$$+ \bar{m}_i^{III} < d_{i,i-1}, R(\theta_{i+1,i})d_{i+1,i+2} > \tag{3.3.8b}$$

$$\tilde{M}_{i,i+2} = -\bar{m}_i^{II} < d_{i,i+1}, R(\theta_{i+2,i})d_{i+2,i+1} >$$

$$- \bar{m}_{i+1}^{III} < d_{i,i+1}, R(\theta_{i+2,i})d_{i+2,i-1} >$$

$$- \bar{m}_{i+2}^{II} < d_{i,i-1}, R(\theta_{i+2,i})d_{i+2,i-1} >$$

$$- \bar{m}_i^{III} < d_{i,i-1}, R(\theta_{i+2,i})d_{i+2,i+1} > \tag{3.3.8c}$$

for $i = 0, 1, 2, 3 \pmod 4$.

Remark 3.3.1: Let the four bars be labeled sequentially such that $l_1 = \bar{s}$. It is clear that the dependence of T on θ_i, $i = 0, 1, 2, 3$ can be rearranged so that kinetic energy T depends on θ_{10} and θ_{32}. In addition, from (3.1.2) or (3.2.4), θ_i, $i = 0, 1, 2, 3$ depend

on each other by constraint. From [36], we have the following explicit dependence of θ_2 and θ_3 on θ_0 and θ_1:

$$\theta_2 = \theta_0 + f(\theta_{10}; sgn(sin(\theta_{32})))$$

$$\theta_3 = \theta_0 + g(\theta_{10}; sgn(sin(\theta_{32}))),$$

where

$$f(\theta_{10}; sgn(sin(\theta_{32}))) = \begin{cases} \pi - \beta^+ & \text{if } sin(\theta_{32}) > 0; \\ \pi - \beta^- & \text{if } sin(\theta_{32}) < 0. \end{cases}$$

for

$$\beta^\pm = arctan_2(-l_1 sin(\theta_{10}), l_0 + l_1 cos(\theta_{10}))$$

$$\pm arctan_2((4l_3^2 l_2^2 - (l_0^2 + l_1^2 - l_3^2 - l_2^2 + 2l_0 l_1 cos(\theta_{10}))^{\frac{1}{2}},$$

$$2l_3(l_0^2 + l_1^2 + 2l_0 l_1 cos(\theta_{10}))^{\frac{1}{2}})$$

and

$$g(\theta_{10}; sgn(sin(\theta_{32}))) = \begin{cases} -\pi + \gamma^+ & \text{if } sin(\theta_{32}) > 0; \\ -\pi + \gamma^- & \text{if } sin(\theta_{32}) < 0. \end{cases}$$

for

$$\gamma^\pm = arctan_2(l_3 sin(\beta^\pm) + l_1 sin(\theta_{10}), l_0 - l_3 cos(\beta^\pm) + l_1 cos(\theta_{10})).$$

In above expression $arctan_2$ is defined as follows. Let

$$x = cos(\theta) \quad \text{and} \quad y = sin(\theta)$$

$$\theta_1 = tan^{-1} \frac{y}{x} \quad (principle \ value).$$

Then,

$$\theta = arctan_2(y, x) \triangleq \begin{cases} \theta_1, & \text{if } x > 0; \\ \theta_1 + sgn(y)\pi, & \text{if } x < 0; \\ sgn(y)\frac{\pi}{2}, & \text{if } x=0. \end{cases}$$

From the above expressions we see that the value of θ_{32} depends not only on θ_{10}, but also on $sgn(sin(\theta_{32}))$, i.e., leading or lagging form. For a Grashof linkage, on each component of the configuration space, the sign of $sin(\theta_{32})$ does not change. For a non-Grashof linkage, we note that for almost all values of θ_{32} there exists a neighborhood about θ_{32} such that θ_{32} depends on θ_{10} uniquely, except at the dead points, where $sin(\theta_{32}) = 0$. As we have mentioned in Remark 3.2.8(3), to deal with this problem in such neighborhoods, an easy way is to re-label four bars such that $sin(\theta_{32}) \neq 0$. One

can always do so since all four bars cannot be aligned in this case. Based on the above discussion we conclude that, in general, at each point in Q, every element of the matrix $\tilde{M}$ can be expressed as a function of θ_{10}. ∎

Differentiating the loop constraints (3.1.2), one can get a relation between (ω_0, ω_1) and (ω_2, ω_3):

$$\begin{pmatrix} \omega_2 \\ \omega_3 \end{pmatrix} = N \begin{pmatrix} \omega_0 \\ \omega_1 \end{pmatrix}, \tag{3.3.9}$$

where

$$N = \begin{pmatrix} -\frac{l_0 \sin(\theta_3 - \theta_0)}{l_2 \sin(\theta_3 - \theta_2)} & -\frac{l_1 \sin(\theta_3 - \theta_1)}{l_2 \sin(\theta_3 - \theta_2)} \\ \frac{l_0 \sin(\theta_2 - \theta_0)}{l_3 \sin(\theta_3 - \theta_2)} & \frac{l_1 \sin(\theta_2 - \theta_1)}{l_3 \sin(\theta_3 - \theta_2)} \end{pmatrix}. \tag{3.3.10}$$

Again, following the same line of reasoning as in Remark 3.3.1, the matrix N is well-defined locally in general.

We summarize the above discussion in the following theorem.

Theorem 3.3.2: The kinetic energy of a floating four-bar linkage can be represented as

$$T = \frac{1}{2} < \omega, M\omega > + \frac{1}{2} m \|\dot{r}_c\|^2, \tag{3.3.11}$$

where $\omega = (\omega_0, \omega_1)^T$ and

$$M = (I \quad N^T) \tilde{M} \begin{pmatrix} I \\ N \end{pmatrix} \tag{3.3.12}$$

for matrix $\tilde{M}$ given in (3.3.8) and matrix N given in (3.3.10). The elements of matrix M are, locally, functions of θ_{10}. ∎

Remark 3.3.3: As in the case of planar multibody systems with open chains, the kinetic energy of a four-bar linkage depends on the relative angles of the bars, although such dependence appears more complicated. By capturing this property, we are able to study symmetry and reduction, as will be seen in next section. ∎

Before ending this section, we give a property of the matrix N which will be used in next section.

Proposition 3.3.4: If N is well defined,

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = N \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

43

Proof: Premultiplying (3.1.2) by $R(-\theta_2)$ and $R(-\theta_3)$, we get

$$l_0 sin(\theta_0 - \theta_2) + l_1 sin(\theta_1 - \theta_2) + l_3 sin(\theta_3 - \theta_2) = 0$$

and

$$l_0 sin(\theta_0 - \theta_3) + l_1 sin(\theta_1 - \theta_3) + l_2 sin(\theta_2 - \theta_3) = 0$$

respectively. The result follows immediately.
∎

3.4 Symmetry, Integral and Reduction

We shall show here that a floating four-bar linkage is a simple mechanical system with symmetry in the sense of Smale, which has been defined in Section 2.1. For simplification, we make two assumptions. First, we restrict attention to the type of Grashof linkages, i.e., Grashof condition $\bar{s} + \bar{l} < \bar{p} + \bar{q}$ holds, and assume that the linkage has been in either leading or lagging form. The four bars are labeled sequentially with $l_1 = \bar{s}$. For the non-Grashof case we will give a discussion in the final remark of this section. Secondly, we assume that the inertial observer is placed at the center of mass of the system. We are able to do so since the kinetic energy in (3.2.11) is invariant under the translation in inertial space. In [44] this process is explained via *symplectic reduction* by the translation group $\Re^2$.

Keeping the above assumptions in mind, we can identify each ingredient in the 4-tuple (Q, K, V, G) of simple mechanical system with symmetry. The configuration space now is simply

$$Q = S^1 \times S^1.$$

Each point q in Q is parametrized by $q = (\theta_0, \theta_1)$, the absolute angles of 0-th and 1-st bar. The Riemannian metric on Q is, for each $q \in Q$,

$$K(q)(v_q, w_q) = < v_q, M(q)w_q >,$$

where $v_q, w_q \in T_q Q \simeq \Re^2$ and matrix M is given in Theorem 3.3.2. The potential energy V has been assumed to be zero. Let $G = S^1 \simeq SO(2)$ be the Lie symmetry group. Its action Φ on Q is defined by, for any $\phi \in G$,

$$\Phi(\phi, (\theta_0, \theta_1)) = (\theta_0 + \phi, \theta_1 + \phi). \tag{3.4.1}$$

44

As have been discussed in Remark 3.3.1 and 3.3.3, matrix M depends on θ_{10} only. This implies that G acts on Q by isometries. Therefore, we conclude that a floating four-bar linkage is a simple mechanical system with symmetry. In addition, for this system, the associated Lagrangian is given by

$$L(q, v_q) = \frac{1}{2} K(q)(v_q, v_q) = \frac{1}{2} < v_q, \mathbf{M}(\theta_{10})v_q >, \qquad (3.4.2)$$

where $v_q = (\dot{\theta}_0, \dot{\theta}_1)^T$, and the Hamiltonian is given by

$$H(q, p) = \frac{1}{2} < p, \mathbf{M}^{-1}(\theta_{10})p >, \qquad (3.4.3)$$

where $p = (p_0, p_1)^T \stackrel{\triangle}{=} K^b(q)(v_q) = \mathbf{M}(\theta_{10})v_q \in T_q^*Q \simeq \Re^2$. It is clear that H is G-invariant.

Remark 3.4.1: From the above discussion, we see that the dynamic structure of a floating four-bar linkage is the same as that of a coupled, planar two-body [43] although the explicit forms of the Riemannian metrics are very different. Therefore, in some essential ways, the description of these two types of systems are parallel to each other. Because of this, we will not prove the following assertions about reduction in detail. For reference, see [43]. ∎

Let $\xi = \bar{\xi}$, a constant in $\mathcal{G} = \Re$ (the Lie algebra of S^1). One can show that the infinitesimal generator of the action Φ given in (3.4.1) corresponding to ξ is simply

$$\xi_Q(q) = \bar{\xi}(1, \ 1)^T. \qquad (3.4.4)$$

Then, the momentum map $\mathbf{J} : T^*Q \rightarrow \mathcal{G} = so^*(2)$ for the cotangent lift action of Φ (cf. Section 2.1) determined by (2.1.26) is

$$\langle \mathbf{J}(q, p), \xi \rangle = < p, \xi_Q(q) > = \bar{\xi}(p_0 + p_1) \stackrel{\triangle}{=} \nu. \qquad (3.4.5)$$

From Noether's Theorem, it is easy to show that ν or $\mu \stackrel{\triangle}{=} \mathbf{J}(q, p) = p_0 + p_1$ is conserved along trajectories of X_H for the Hamiltonian H in (3.4.3) and that μ is simply the net angular momentum of the floating four bar linkage relative to an observer at the system center of mass.

Since the dynamical trajectories are confined to a level set of the form $\mathbf{J}^{-1}(\mu)$ and the group S^1, viewed as the isotropy subgroup of the momentum value μ, acts freely

45

on $\mathbf{J}^{-1}(\mu)$, one gets the symplectically reduced dynamics X_{H_μ} on the reduced phase space $P_\mu = \mathbf{J}^{-1}(\mu)/S^1 \simeq S^1 \times \Re^1$, where H_μ is the reduced Hamiltonian. For details of symplectic reduction, see [1].

As in [44] it is also possible to Poisson-reduce the dynamics. Following the standard Poisson reduction framework which has been shown in Section 2.1, we have the following results.

Since the manifold $M = T^*(S^1 \times S^1)$ parameterized by $z = (\theta_0, \theta_1, p_0, p_1)$ is symplectic, a Poisson structure $\{\ ,\ \}_M$ can be constructed, as we indicated in Remark 2.1.1. It is easy to show that this structure is of the form

$$\{f_1, f_2\}_M(z) = \sum_{i=0}^{1} \left(\frac{\partial f_1}{\partial \theta_i} \cdot \frac{\partial f_2}{\partial p_i} - \frac{\partial f_1}{\partial p_i} \cdot \frac{\partial f_2}{\partial \theta_i} \right) \tag{3.4.6}$$

for all f_1, $f_2 \in C^\infty(M)$. The action of $G = S^1$ on Q given by (3.4.7) is free and proper. The quotient $P = T^*(S^1 \times S^1)/S^1 \simeq S^1 \times \Re^2$ carries a reduced Poisson structure. Parameterizing P by $\tilde{z} = (\theta_{10}, p_0, p_1)$, the noncanonical Poisson bracket on P is given by,

$$\begin{aligned}
\{\tilde{f}_1, \tilde{f}_2\}_P(\tilde{z}) &= \frac{\partial \tilde{f}_1}{\partial \theta_{10}} \cdot \left(\frac{\partial \tilde{f}_2}{\partial p_1} - \frac{\partial \tilde{f}_2}{\partial p_0} \right) - \frac{\partial \tilde{f}_2}{\partial \theta_{10}} \cdot \left(\frac{\partial \tilde{f}_1}{\partial p_1} - \frac{\partial \tilde{f}_1}{\partial p_0} \right) \\
&= \nabla \tilde{f}_1^T \cdot \Lambda \nabla \tilde{f}_2,
\end{aligned} \tag{3.4.7}$$

where $\tilde{f}_1, \tilde{f}_2 \in C^\infty(P)$, $\nabla \tilde{f}_i = \left(\frac{\partial \tilde{f}_i}{\partial \theta_{10}}\ \frac{\partial \tilde{f}_i}{\partial p_0}\ \frac{\partial \tilde{f}_i}{\partial p_1} \right)^T$ and $\Lambda = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$. The reduced Hamiltonian $\tilde{H}$ is given by

$$\tilde{H}(\theta_{10}, p_0, p_1) = H(\theta_0, \theta_1, p_0, p_1), \tag{3.4.8}$$

since the matrix $\mathbf{M}$ in (3.3.12) is a matrix of functions of θ_{10}. The reduced dynamics is then immediately given by reduced Hamiltonian vector field

$$X_{\tilde{H}} = \Lambda \nabla \tilde{H}$$

or, explicitly, by differential equations

$$\begin{cases}
\dot{p}_0 = \dfrac{\partial \tilde{H}}{\partial \theta_{10}} \\[2mm]
\dot{p}_1 = -\dfrac{\partial \tilde{H}}{\partial \theta_{10}} \\[2mm]
\dot{\theta}_{10} = \dfrac{\partial \tilde{H}}{\partial p_1} - \dfrac{\partial \tilde{H}}{\partial p_0}.
\end{cases} \tag{3.4.9}$$

Unlike coupled two-body problem, Equation (3.4.9) involves complicated analytic expressions resulting from the substitutions for θ_3 and θ_2 in terms of θ_0 and θ_1 (cf. Remark 3.3.1). Certain qualitative aspects of the reduced dynamics can still be explored such as relative equilibria which we will investigate in the following section.

Remark 3.4.2: For non-Grashof mechanisms, following Remark 3.3.1, the system still has symmetry group $G = S^1$. It is clear that, in this case, if θ_0 and θ_1 are just parameters of configuration space $S^1 \times S^1$, instead of real angles of 0-th and 1-st bars, all the results in this section are valid. However, for convenience, in next section we will still use physical angles θ_0 and θ_1 for non-Grashof mechanisms. Then, θ_0 and θ_1 will play the role of *local coordinates* for $S^1 \times S^1$ and, consequently, (3.4.5), (3.4.7) and (3.4.9) will be the expressions of momentum, reduced Poisson bracket and reduced dynamics in local coordinates, respectively. ∎

One of the advantages of applying reduction theory to mechanical systems is that it helps to make the dynamics of the system more transparent. For our system which is of four dimensions in phase space, the reduction process make it possible to display the dynamics by phase portraits on a lower dimensional space. To illustrate this, we show two examples here.

Example 3.4.3:

Consider a floating four-bar linkage whose parameters satisfy the Grashof condition and which is of lagging form. In particular, the parameters are chosen as follows.

$$m_0 = 1, \quad m_1 = 1, \quad m_2 = 1, \quad m_3 = 1;$$

$$I_0 = 1, \quad I_1 = 1, \quad I_2 = 1, \quad I_3 = 1;$$

$$\mathbf{d}_{03} = (-1.5, 1), \quad \mathbf{d}_{01} = (1.5, 1), \quad \mathbf{d}_{10} = (-0.5, 1.3), \quad \mathbf{d}_{12} = (0.5, 1.3),$$

$$\mathbf{d}_{21} = (-1.5, 1), \quad \mathbf{d}_{23} = (1.5, 1), \quad \mathbf{d}_{32} = (-2, \lambda), \quad \mathbf{d}_{30} = (2, \lambda).$$

Following the same procedure as in [44], the dynamics can be further reduced to a symplectic leaf. The Hamiltonian on a leaf is a function of θ_{10} and $\rho \triangleq \frac{1}{2}(p_1 - p_0)$ and the dynamics on the leaf is given by
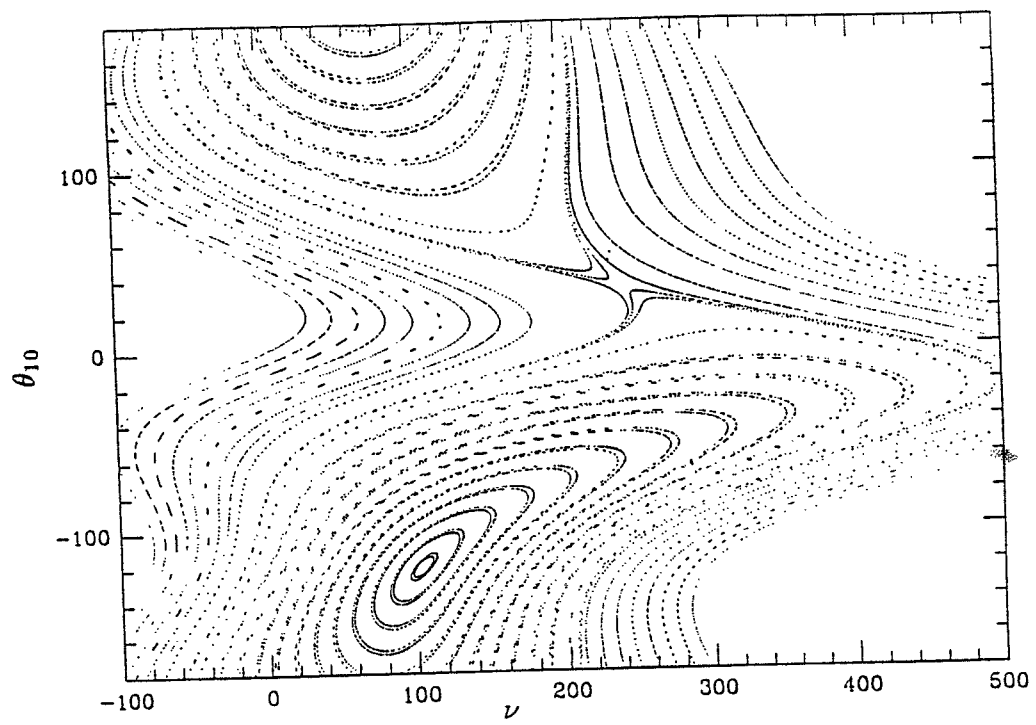
(a) $\lambda = 0$
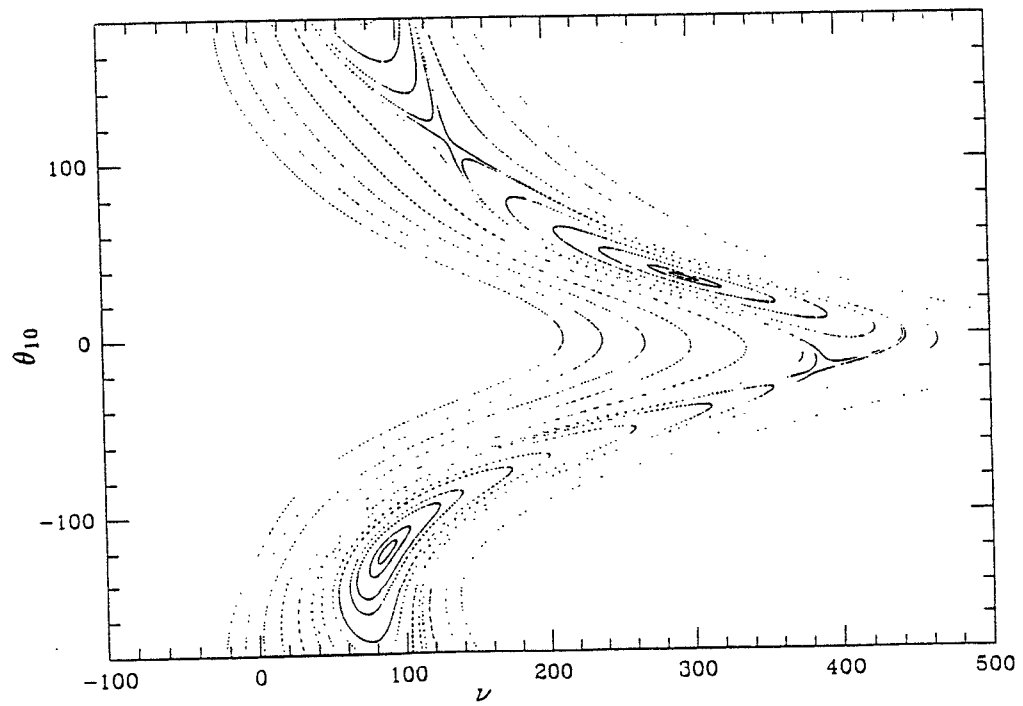
(b) $\lambda = -5$

Figure 3.4.1 Phase portraits for a Grashof linkage

$$\frac{d\theta_{10}}{dt} = \frac{\partial H}{\partial \rho}(\theta_{10}, \rho), \qquad \frac{d\rho}{dt} = -\frac{\partial H}{\partial \theta_{10}}(\theta_{10}, \rho)$$

Therefore, for fixed value of angular momentum, one can draw the phase portrait on a cylinder with coordinates θ_{10} and ρ. Figure 3.4.1 shows the phase portraits for $\lambda = 0$ and $\lambda = -5$, respectively. On Figure 3.4.1(a), we see there is one center and one saddle point. However, in Figure 3.4.1(b), there are two centers and two saddles. The difference is caused by an offset in the position of center of mass of 3rd bar from the line connecting the two joints on that bar. ∎

Example 3.4.4:

Consider a floating four-bar linkage whose parameters satisfy the non-Grashof condition and the parameters are chosen as follows.

$$m_0 = 1, \quad m_1 = 1, \quad m_2 = 1, \quad m_3 = 1;$$

$$I_0 = 1, \quad I_1 = 1, \quad I_2 = 1, \quad I_3 = 1;$$

$$\mathbf{d}_{03} = (-2, \lambda), \qquad \mathbf{d}_{01} = (2, \lambda), \qquad \mathbf{d}_{10} = (-1.5, -1), \qquad \mathbf{d}_{12} = (1.5, -1),$$

$$\mathbf{d}_{21} = (-1.5, -1.4), \quad \mathbf{d}_{23} = (1.5, -1.4), \quad \mathbf{d}_{32} = (-1.5, -1), \quad \mathbf{d}_{30} = (1.5, -1).$$

Instead of displaying the dynamics on position-momentum phase space, here we try to show it on position-velocity phase space. Since the dynamics cannot be further reduced so that it depends on one relative angle and corresponding angular velocity, one has to display the phase portrait in three dimensional space. For a given angular momentum, the phase portrait sits on a surface in this space. In this example, this space is parameterized by $(\dot{\theta}_{10}, \hat{\theta}, \dot{\theta}_{03})$, where $\hat{\theta}$ is defined in (3.2.6). Figure 3.4.2 shows the phase portraits corresponding to $\lambda = -15$ and $\lambda = 0$, respectively. From Figure 3.4.2, we see, again, the change of numbers of centers and saddles. ∎

In the next section, we will show how to compute the centers and saddle points and associated bifurcations.
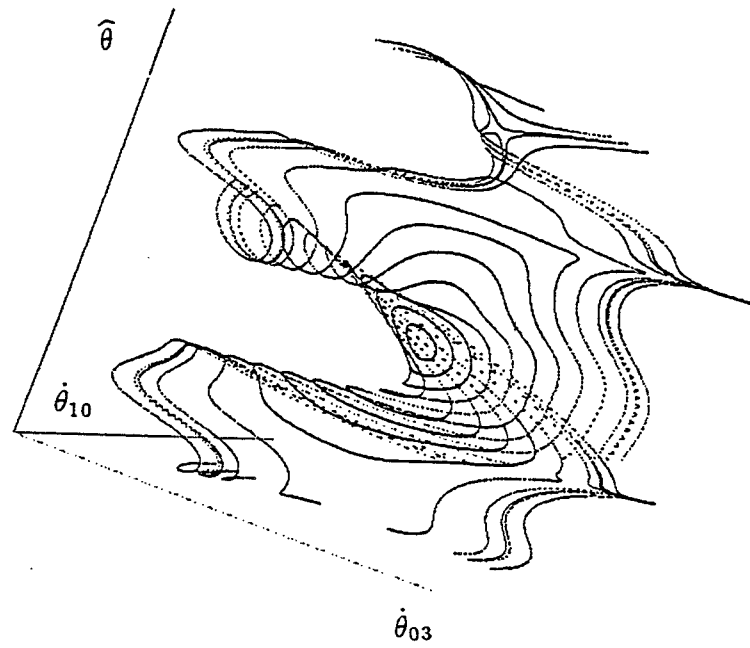
(a) $\lambda = -15$
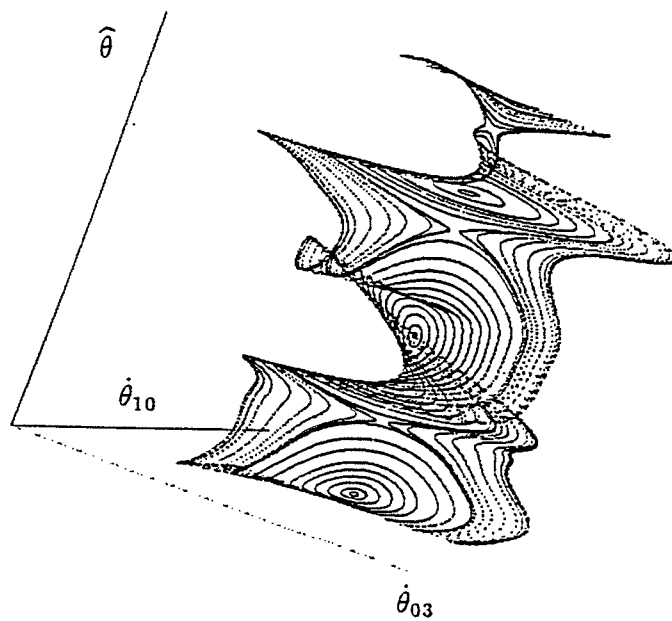
(b) $\lambda = 0$

Figure 3.4.2 Phase portraits for a non-Grashof linkage

3.5 Relative Equilibria and Bifurcations

We first give the definition of relative equilibria and recall Smale's theorem [42].

Consider a Hamiltonian system (M, Ω, X_H), where M is a $2n$-dimensional manifold, Ω is a symplectic two-form on M and, for some smooth Hamiltonian H on M, X_H is Hamiltonian vector field determined by (2.1.9). Let G be a Lie group acting on M symplectically and leaving H invariant. The following definition of a relative equilibrium is standard [1].

Definition 3.5.1: Let $F^t_{X_H}$ be the flow of X_H on M. Then $z_e \in M$ is a relative equilibrium if $F^t_{X_H}(z_e)$ is a stationary motion, i.e., there exists $\xi \in \mathcal{G}$ such that

$$F^t_{X_H}(z_e) = exp(t\xi)(z_e),$$

where $\mathcal{G}$ is the Lie algebra of the group G. ∎

Remarks 3.5.2:

(1) Let $X_{\tilde{H}}$ be the Poisson reduced vector field as shown in (2.1.21). Then z_e is a relative equilibrium if and only if

$$X_{\tilde{H}}(\pi(z_e)) = 0,$$

where $\pi : M \to M/G$ is the canonical projection.

(2) A physical interpretation of relative equilibria is that if the dynamics of a system is rotationally invariant, the dynamical orbit of a relative equilibrium appears to be a fixed point for an observer in a suitable uniformly rotating coordinate system.

∎

Given a simple mechanical system with symmetry, (Q, K, V, G), the associated Hamiltonian system can be constructed by (T^*Q, Ω_0, X_H), where Ω_0 is the canonical two-form and the Hamiltonian H is given by (2.1.25). The symmetry group G acts on T^*Q by lifting (cf. (2.1.3)). The following theorem is due to Smale [42].

Theorem 3.5.3: For a simple mechanical system with symmetry (Q, K, V, G), define *augmented potential function* by

$$V_\xi : Q \to \Re : q \mapsto V(q) - \frac{1}{2}K(q)(\xi_Q(q), \xi_Q(q)) \tag{3.5.1}$$

51

for each $\xi \in \mathcal{G}$, where ξ_Q is infinitesimal generator of the action corresponding to ξ. Then $z_e = (q_e, p_e) \in T^*Q$ is a relative equilibrium if and only if q_e is a critical point of V_ξ for some $\xi \in \mathcal{G}$ and $p_e = K^\flat(q)(\xi_Q(q_e))$.

∎

Remarks 3.5.4:

(1) It can be shown that, for a given $\xi \in \mathcal{G}$, V_ξ has the symmetry,

$$V_\xi(\Phi_g(x)) = V_\xi(x) \tag{3.5.2}$$

for all $g \in G_\xi := \{g \in G | Ad_g\xi = \xi\}$. If the action Φ is free and proper, then the quotient space, Q/G_ξ, is a smooth manifold and $\pi_\xi : Q \to Q/G_\xi$ is a submersion. Thus V_ξ induces a function $\hat{V}_\xi$ on Q/G_ξ such that

$$V_\xi = \hat{V}_\xi \circ \pi_\xi.$$

(2) If G is Abelian, $G_\xi = G$. In this case, we refer to Q/G_ξ as the *shape space*, and the points $\pi_\xi(q_e)$ as the relative equilibrium shapes. It can be shown that, in this case, if $\pi_\xi(q_e)$ is a local minimizer of $\hat{V}_\xi$, the corresponding $z_e = (q_e, p_e)$ is a stable relative equilibrium, and if $\pi_\xi(q_e)$ is a local maximizer of $\hat{V}_\xi$, the corresponding $z_e = (q_e, p_e)$ is an unstable relative equilibrium.

(3) One can also use the *amended potential* in the sense of Smale to study the relative equilibria (see [41]).

∎

Now we address the problem of computing relative equilibria of floating four-bar linkages. In Section 3.4 we have shown that this mechanism is a simple mechanical system with symmetry. So Smale's Theorem 3.5.3 is applicable to the system. Recall that, by letting $\xi = \bar{\xi}$, a constant in $\mathcal{G} = \Re$, the infinitesimal generator of action (3.4.1) with respect to this ξ is $\xi_Q(q) = \bar{\xi}(1,1)^T$. Hence, by Theorem 3.5.3, (q_e, p_e) is a relative equilibrium point on T^*Q if and only if q_e is a critical point of the function

$$V_\xi(\theta_0, \theta_1) = -\bar{\xi}^2(1,1)\mathbf{M}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

where $q_e = ((\theta_0)_e, (\theta_1)_e)$, $p_e = ((p_0)_e, (p_1)_e)$ and matrix $\mathbf{M}$ is defined in (3.3.12). Applying Proposition 3.3.4 and the 4×4 matrix $\tilde{\mathbf{M}}$ defined in (3.3.8), we have

$$V_\xi(\theta_0, \theta_1) = -\bar{\xi}^2 \mathbf{e}^T \tilde{\mathbf{M}} \mathbf{e}, \tag{3.5.3}$$

52

where $\mathbf{e} = (1\ 1\ 1\ 1)^T$. Since the elements of matrix $\tilde{\mathbf{M}}$ are, locally, functions of θ_{10} or constants, the above V_ξ satisfies (3.5.2) for all $g \in S^1$. It follows that the induced function $\hat{V}_\xi$ shown in Remark 3.5.4(1) is

$$\hat{V}_\xi(\theta_{10}) = V_\xi(\theta_0, \theta_1). \qquad (3.5.4)$$

Then, the critical points q_e of V_ξ will make $(\theta_{10})_e \stackrel{\triangle}{=} (\theta_1)_e - (\theta_0)_e$ to be the critical points of $\hat{V}_\xi$. At relative equilibrium, the relative angles between bars keep unchanged and the whole system rotates around the system center of mass with constant angular velocity.

Remarks 3.5.5:

(1) Following Remark 3.3.1, it should be noted that, in general, given the value of $(\theta_{10})_e$ one cannot tell what the relative equilibrium shape of the linkage looks like. The particular form (leading or lagging) has to be indicated at the same time.

(2) $\hat{V}_\xi$ given in 3.5.4 is the *locked inertia* if $\bar{\xi} = 1$. Its value at θ_{10} equals the value of the moment of inertia of the corresponding frozen system, i.e., the system with all joints locked, about its center of mass. The above result shows that a relative equilibrium shape corresponds to a frozen system which has maximum or minimum value of moment of inertia within all possible frozen systems. Since $\bar{\xi}$ is a constant, without loss of generality, we let $\bar{\xi} = 1$ later. ∎

In the rest of this section, we are particularly interested in assemblies which admit configurations with reflection symmetry, which will be called *symmetric configuration*. Applying the notations in Section 3.1, a floating four-bar linkage is of *symmetric type* if, with proper consecutive labeling of the bars,

$$m_1 = m_3, \qquad l_1 = l_3 \qquad (3.5.5\text{a})$$

$$|\mathbf{d}_{01}| = |\mathbf{d}_{03}|, \ |\mathbf{d}_{10}| = |\mathbf{d}_{30}|, \ |\mathbf{d}_{12}| = |\mathbf{d}_{32}|, \ |\mathbf{d}_{21}| = |\mathbf{d}_{23}| \qquad (3.5.5\text{b})$$

and

$$\mathbf{d}_{10}^T \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{d}_{30} > 0. \qquad (3.5.5\text{c})$$

In other words it can form symmetric shapes as shown in Figure 3.5.1. It is not hard to verify that any four-bar linkage which satisfies (3.5.5) has two such symmetric configurations.
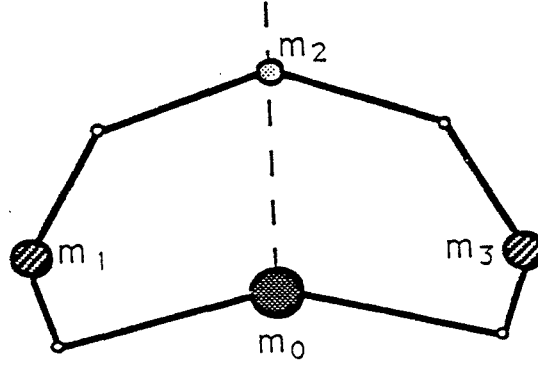
Figure 3.5.1 A symmetric configuration

Although it is not easy to find the critical points of $\hat{V}_\xi$ analytically, for a particular example one can easily find them numerically. Unlike the planar two body problem [43] for which the dimension of the shape space equals 1, here the function $\hat{V}_\xi$ has many parameters. A natural question is to determine how these parameters affect the relative equilibria, e.g. their numbers and location on shape space, etc.. *Of course, it is difficult to answer this question for completely arbitrary choice of parameters.* However, by leaving one or two particular parameters free such that the assembly preserves its symmetric configurations and freezing all other parameters, one still can observe a nontrivial bifurcation phenomenon. To illustrate this we consider an example.

Example 3.5.6: Let us choose the parameters as follows.

$$m_0 = 1, \quad m_1 = 1, \quad m_2 = 1, \quad m_3 = 1;$$

$$\mathbf{d}_{03} = (-2, \lambda_0), \qquad \mathbf{d}_{01} = (2, \lambda_0), \qquad \mathbf{d}_{10} = (-1.5, -1), \qquad \mathbf{d}_{12} = (1.5, -1),$$

$$\mathbf{d}_{21} = (-1.5, -1.4), \quad \mathbf{d}_{23} = (1.5, -1.4), \quad \mathbf{d}_{32} = (-1.5, -1), \quad \mathbf{d}_{30} = (1.5, -1).$$

Now the assembly has *non-Grashof* structure.

Using $\hat{V}_\xi$, for any λ_0 one can find relative equilibria $(\theta_{10})_e$ for both leading form and lagging form, and hence, corresponding $\hat{\theta}$ which is defined in (3.2.6). As λ_0 varies from $-\infty$ to $+\infty$, one can plot a diagram for $\hat{\theta}$ versus λ_0. Figure 3.5.2 shows the result, in which solid dots represent stable relative equilibria, small circles represent unstable relative equilibria. ∎
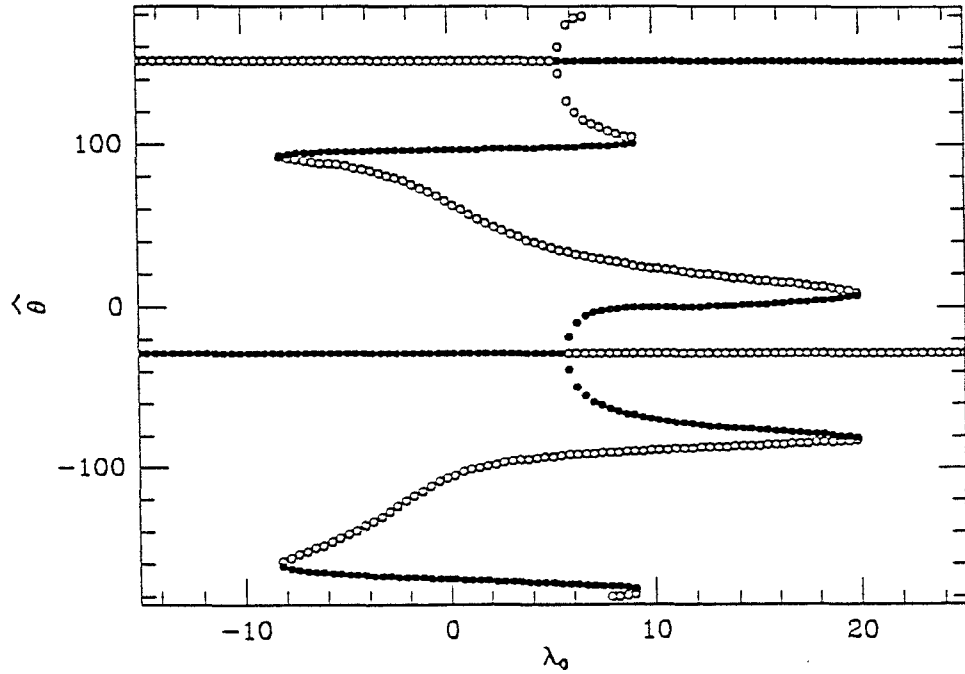
Figure 3.5.2 Bifurcation diagram: an example

From this example one can make the following empirical observations:

(1) There are two unbounded symmetric branches on the diagram and and these branches are bifurcated at some points. The bifurcations appear to be pitchfork bifurcations.

(2) Almost any value of $\hat{\theta}$ can be a relative equilibrium for a particular λ_0. In other words the bifurcation diagram is connected globally.

(3) The number of relative equilibria can be two, six and ten.

The first observation, which relates to the local bifurcation problem, is what we

will concentrate on in the rest of this section. The others will be discussed later.

As we have seen, the function $\hat{V}_\xi$ of a four-bar linkage is a multiple parameter function. One might expect very complicated bifurcation features with respect to these parameters. Here, instead of considering a general structure, we study a special assembly which admits a symmetric configuration. To avoid too many tedious calculations we particularly choose the parameters of the assembly as follows.

$$m_0 = m_1 = m_2 = m_3 = 1; \tag{3.5.6a}$$

$$\mathbf{d}_{03} = (-d_0, \lambda_0), \quad \mathbf{d}_{01} = (d_0, \lambda_0), \quad \mathbf{d}_{10} = (-1, 0), \quad \mathbf{d}_{12} = (1, 0), \tag{3.5.6b}$$

$$\mathbf{d}_{21} = (-d_2, \lambda_2), \quad \mathbf{d}_{23} = (d_2, \lambda_2), \quad \mathbf{d}_{32} = (-1, 0), \quad \mathbf{d}_{30} = (1, 0), \tag{3.5.6c}$$

where d_0 and d_2 are fixed and $d_0 > d_2 > 0$, $\lambda_0, \lambda_2 \in \Re$. Moreover, we consider the non-Grashof case only, i.e., $\bar{s} + \bar{l} > \bar{p} + \bar{q}$.

Figure 3.5.3 shows two symmetric configurations for the above choice of parameters. We will see that although only two parameters λ_0 and λ_2 are left to be free, the bifurcation features with respect to these parameters are still informative.

The function $\hat{V}_\xi$ now has the following form:

$$\hat{V}_\xi = \frac{1}{4}(2\lambda_2 sin(\theta_{32}) + d_2 cos(\theta_{32}) - cos(\theta_{31}) + d_0 cos(\theta_{30})$$

$$+ 2\lambda_2 sin(\theta_{21}) + d_2 cos(\theta_{21}) - d_0 d_2 cos(\theta_{20}) + d_0 cos(\theta_{10}))$$

$$+ \frac{\lambda_0}{4}(2sin(\theta_{30}) + \lambda_2 cos(\theta_{20}) - 2sin(\theta_{10})) + C, \tag{3.5.7}$$

where C is a constant determined by $\mathbf{d}_{ij}$, m_i and the moments of inertia of the bodies, I_i. For $i, j = 0, 1, 2, 3$, $\theta_{ij} = \theta_i - \theta_j$ satisfy the constraint equations

$$d_0 + cos(\theta_{10}) + d_2 cos(\theta_{20}) + cos(\theta_{30}) = 0 \tag{3.5.8a}$$

$$sin(\theta_{10}) + d_2 sin(\theta_{20}) + sin(\theta_{30}) = 0. \tag{3.5.8b}$$

In the following, at symmetric configuration, the variables will be denoted by superscript "s" (say, θ_{10}^s), the formulas will be denoted by "$|_s$" (say, $f(\theta_{10})|_s$). As shown in the example, the bifurcation diagram of relative equilibria will be parameterized by $(\hat{\theta}, \lambda_0)$.
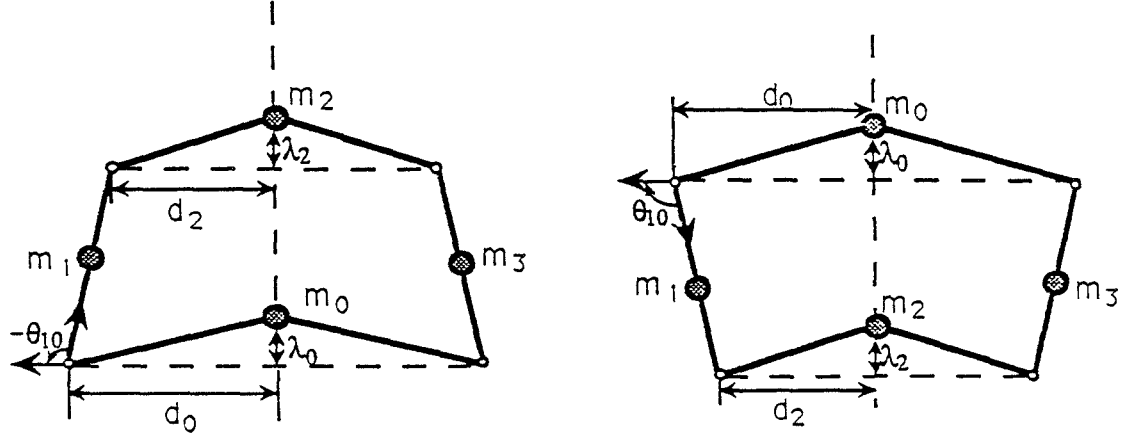
Figure 3.5.3 Symmetric configurations

Theorem 3.5.7: For a floating four-bar linkage with parameters shown in (3.5.6), the bifurcation diagram of relative equilibria has the following properties:

(1) There are two infinite branches in the diagram, one corresponds to θ_{10}^s in the leading form, another one corresponds to θ_{10}^s in the lagging form. We refer to these as the *symmetric branches*.

(2) There exists a constant λ_2^* such that no bifurcation occurs on the symmetric branch of leading form if $\lambda_2 = \lambda_2^*$; and, no bifurcation occurs on the symmetric branch of lagging form if $\lambda_2 = -\lambda_2^*$.

(3) On the symmetric branch of leading (lagging) form, if $\lambda_2 < \lambda_2^* \ (-\lambda_2^*)$, there exists a constant $c_1 \ (c_3)$ such that the relative equilibria are stable for $\lambda_0 < c_1 \ (c_3)$, unstable for $\lambda_0 > c_1 \ (c_3)$, bifurcated for $\lambda_0 = c_1 \ (c_3)$; one the other hand, if $\lambda_2 > \lambda_2^* \ (-\lambda_2^*)$, there exists a constant $c_2 \ (c_4)$ such that the relative equilibria are unstable for $\lambda_0 < c_2 \ (c_4)$, stable for $\lambda_0 > c_2 \ (c_4)$, bifurcated for $\lambda_0 = c_1 \ (c_4)$.

(4) Assume $\lambda_2 \neq \pm\lambda_2^*$. Let

$$\epsilon^\pm = sgn\frac{\pm\epsilon_1\lambda_2^2 + \epsilon_2\lambda_2 \pm \epsilon_3}{\epsilon_4\lambda_2 \pm \epsilon_5} \qquad (3.5.9)$$

and

$$\delta^\pm = sgn(\delta_1\lambda_2 \pm \delta_2), \qquad (3.5.10)$$

where ϵ_i and δ_i are constants which are determined by d_0 and d_2. Then, the bifurcation on the symmetric branch of leading form will be supercritical pitchfork if $\epsilon^+\delta^+ < 0$. It will be subcritical pitchfork if $\epsilon^+\delta^+ > 0$. Similarly, the bifurcation on the symmetric branch of lagging form will be supercritical pitchfork if $\epsilon^-\delta^- < 0$. It will be subcritical pitchfork if $\epsilon^-\delta^- > 0$.

Remarks 3.5.8:

(1) Based on the techniques of bifurcation theory shown in Appendix 3.A, the proof of the above assertions is elementary. However, two aspects have to be considered before the bifurcation theory is applied. First, one has to determine the *existence* of a bifurcation and where the branches are bifurcated. Secondly, applying the techniques of bifurcation theory shown in Appendix 3.A will involve as high as fourth order derivatives of the function $\hat{V}_\xi$ with respect to θ_{10}. However, as we have shown in Remark 3.3.1, matrix M depends on θ_{10} in a very complicated way which makes explicit representing of high order derivative of $\hat{V}_\xi$ to θ_{10} almost impossible. Therefore, when applying the techniques of bifurcation theory to this particular system, one should consider the closure constraint equation (3.5.8) simultaneously.

(2) Although the proof of the above assertions is elementary, it requires a large effort in calculations. We used MACSYMA to handle these computations. In the following, we only give a sketch of the proof. ∎

Proof of Theorem 3.5.7: Note that the function $\hat{V}_\xi$ can be written as a function of relative angles θ_{10}, θ_{20}, and θ_{30}, which are related through constraint equations (3.5.8). From (3.5.8), we can consider (locally) θ_{20} and θ_{30} as the functions of θ_{10}. Again, from (3.5.8) one can generate the quantities $\frac{\partial^i \theta_{20}}{\partial \theta_{10}^i}|_s$ and $\frac{\partial^i \theta_{30}}{\partial \theta_{10}^i}|_s$ for any positive integer i. Moreover, from Figure 3.5.3 it is easy to see that at symmetric configuration

$$\theta_{20}^s = \pi \qquad \text{and} \qquad \theta_{10}^s = -\theta_{30}^s. \qquad (3.5.11)$$

With above considerations, one can have closed form expressions for $\frac{\partial^i \theta_{20}}{\partial \theta_{10}^i}|_s$ and $\frac{\partial^i \theta_{30}}{\partial \theta_{10}^i}|_s$. For instance,

$$\begin{cases} \frac{\partial \theta_{20}}{\partial \theta_{10}}|_s = \frac{2}{d_2} cos(\theta_{10}^s), \\ \frac{\partial \theta_{30}}{\partial \theta_{10}}|_s = 1 \end{cases} \tag{3.5.12}$$

and

$$\begin{cases} \frac{\partial^2 \theta_{20}}{\partial \theta_{10}^2}|_s = \frac{2cos^2(\theta_{10}^s)}{d_2^2 sin(\theta_{10}^s)}(d_2 - 2cos(\theta_{10}^s)), \\ \frac{\partial^2 \theta_{30}}{\partial \theta_{10}^2}|_s = \frac{2cos(\theta_{10}^s)}{d_2 sin(\theta_{10}^s)}(d_2 - 2cos(\theta_{10}^s)) \end{cases} \tag{3.5.13}$$

and so on.

To prove assertion (1) in the statement of the theorem, one needs to show that at symmetric configuration, the equation

$$\frac{\partial \hat{V}_\xi}{\partial \theta_{10}}|_s = 0$$

does not depend on λ_0. Concentrating on the term involving λ_0 in $\hat{V}_\xi$ and applying (3.5.11) and (3.5.12), one can show that the first derivative of that term with respect to θ_{10} at symmetric configuration is zero. Since the rest of the terms of $\frac{\partial \hat{V}_\xi}{\partial \theta_{10}}|_s = 0$ are still functions of θ_{10}^s, one can see two infinite symmetric branches in the bifurcation diagram for two different θ_{10}^s. Assertion (1) is thus proved.

Applying (3.5.12) and (3.5.13), one can show that the second derivative of the function $\hat{V}_\xi$ at symmetric configuration has the form

$$\frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10}^2}|_s = \frac{\partial^2 \Pi}{\partial \theta_{10}^2}|_s - \lambda_0 \left(\frac{d_2 - 2cos^3(\theta_{10}^s)}{d_2 sin(\theta_{10}^s)} + \lambda_2 \frac{cos^2(\theta_{10}^s)}{d_2^2} \right), \tag{3.5.14}$$

where Π is the summation of the terms not involving λ_0 in $\hat{V}_\xi$. It is obvious that when

$$\lambda_2 = \lambda_2^* \triangleq \frac{2d_2 cos^2(\theta_{10}^s) - d_2^2}{cos^2(\theta_{10}^s) sin(\theta_{10}^s)}, \tag{3.5.15}$$

$\frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10}^2}|_s$ will not depend on λ_0. One can also show that with (3.5.15), $\frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10}^2}|_s \neq 0$ under assumptions (3.5.6) and (3.5.7). This means that bifurcation may not occur on either symmetric branch of leading form or symmetric branch of lagging form. Note that on these different forms $cos(\theta_{10}^s)$ has the same value, $sin(\theta_{10}^s)$ has the same absolute value but different sign. (See Figure 3.5.3) Thus, assertion (2) is proved.

As we have known earlier, the stability of relative equilibria depends on the sign of $\frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10}^2}|_s$. From (3.5.14) we see that $\frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10}^2}|_s$ is a linear function of λ_0. Using the λ_2^* in (3.5.15), the proof of (3) is straight forward.

To prove assertion (4), we apply the bifurcation theory mentioned in Appendix 3.A. Let λ_0^* denote c_i in assertion (3) for some suitable i. One can show that at $(\theta_{10}, \lambda_0^*)$,

$$\frac{\partial \hat{V}_\xi}{\partial \theta_{10}} = \frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10}^2} = \frac{\partial^3 \hat{V}_\xi}{\partial \theta_{10}^3} = \frac{\partial^2 \hat{V}_\xi}{\partial \theta_{10} \partial \lambda} = 0. \tag{3.5.16}$$

Moreover

$$\frac{\partial^4 \hat{V}_\xi}{\partial \theta_{10}^4}(\theta_{10}^s, \lambda_0^*) = \frac{\pm \epsilon_1 \lambda_2^2 + \epsilon_2 \lambda_2 \pm \epsilon_3}{\epsilon_4 \lambda_2 \pm \epsilon_5} \tag{3.5.17}$$

and

$$\frac{\partial^3 \hat{V}_\xi}{\partial \theta_{10}^2 \partial \lambda_0}(\theta_{10}^s, \lambda_0^*) = \delta_1 \lambda_2 \pm \delta_2, \tag{3.5.18}$$

where " $+$ " corresponds to leading form, " $-$ " corresponds to lagging form and

$$\epsilon_1 = -12 d_0^2 (d_2 - d_0)^2 (d_2^3 - d_0 d_2^2 - d_0^2 d_2 - 2d_2 + d_0^3 - 4d_0);$$

$$\epsilon_2 = 48 d_0^2 d_2 ((d_2 - d_0)^2 - 1)(d_2^2 - d_0^2) \sqrt{4 - (d_0 - d_2)^2};$$

$$\epsilon_3 = 48 d_0^2 d_2^2 (d_2 - d_0)^2 (d_2^3 - d_0 d_2^2 - d_0^2 d_2 - 4d_2 + d_0^3 - 2d_0);$$

$$\epsilon_4 = d_2^3 ((d_2 - d_0)^2 - 4)(d_2 - d_0)^2 \sqrt{4 - (d_0 - d_2)^2};$$

$$\epsilon_5 = 2 d_2^4 ((d_2 - d_0)^2 - 4)((d_2 - d_0)^3 - 4d_2);$$

$$\delta_1 = -\frac{(d_2 - d_0)^2}{4 d_2^2};$$

$$\delta_2 = -\frac{(d_2^3 + 3 d_0 d_2^2 - 3 d_0^2 d_2 + 4d_2 + d_0^3)}{2 d_2 \sqrt{4 - (d_0 - d_2)^2}}.$$

Since (3.5.17) and (3.5.18) are not zero in general, applying the Lemma 3.A.3 in Appendix 3.A, we can say $\frac{\partial \hat{V}_\xi}{\partial \theta_{10}}$ is strongly equivalent to the normal form of pitchfork bifurcation. In addition, the type of pitchfork bifurcation depends on the sign of (3.5.17) and (3.5.18). The assertion (4) is proved. ∎

Remarks 3.5.9:

(1) The condition of $\lambda_2 \neq \lambda_2^*$ guarantees that (3.5.18) and the denominator of (3.5.17) are not zero.

(2) In general the bifurcation changes from a supercritical one to subcritical one at the roots of numerator of (3.5.17). ∎

Example 3.5.10: To see how λ_2 changes the bifurcation diagram with parameter λ_0 we give following example. We will concentrate on the symmetric branch with respect

to leading form. Let $d_0 = 2$ and $d_2 = 1$. Then ϵ and δ have the following form

$$\epsilon^+ = sgn\frac{\lambda_2^2 + 6.938\lambda_2 - 3.623}{1.873 - 0.051\lambda_2}$$

$$\delta^+ = sgn(1.025 - 0.028\lambda_2)$$

Then

$$\epsilon^+\delta^+ = \begin{cases} -1, & \text{if } -7.426 < \lambda_2 < 0.488; \\ +1, & \text{otherwise.} \end{cases}$$

Note that the region for λ_2 is an approximation. So we can say that, when $\lambda_2 \in (-7.426, 0.488)$, the pitchfork bifurcation is supercritical. Otherwise, it is subcritical. Figure 3.5.4 shows this result. ∎

Before closing this section, we would like to make a few additional remarks.

Remarks 3.5.11:

(1) Although our discussion only concentrated on a structure of non-Grashof type, a version of Theorem 3.5.7 also holds for the Grashof case.

(2) Up to now we have understood the phenomenon of bifurcation on the symmetric branches. The global analysis of the bifurcations involves massive symbolic computations. However, as shown in the Example 3.5.6 in this section and other simulations, one can numerically determine a global bifurcation diagram. A large body of such numerical simulations show that the branches in the bifurcation diagram are connected. In other words, for any point in shape space, there is a finite λ_0 for which that shape determines a relative equilibrium. This is consistent with the linearity of $\hat{V}_\xi$ in λ_0. This property provides a possibility to control the attitude of a space structure with a closed kinematic chain by simply changing the position of the centers of mass of some bars.

(3) Our results in this paper rely on some ideal conditions, for instance, the symmetry condition (3.5.5), and the absence of external and internal disturbances. One may ask what will happen when these conditions are violated. The answer to this question may relate to the notion of *universal unfolding* in bifurcation theory. Numerical results show that it is possible to use the unfolding property to control the shape of the structure near the bifurcation point. ∎
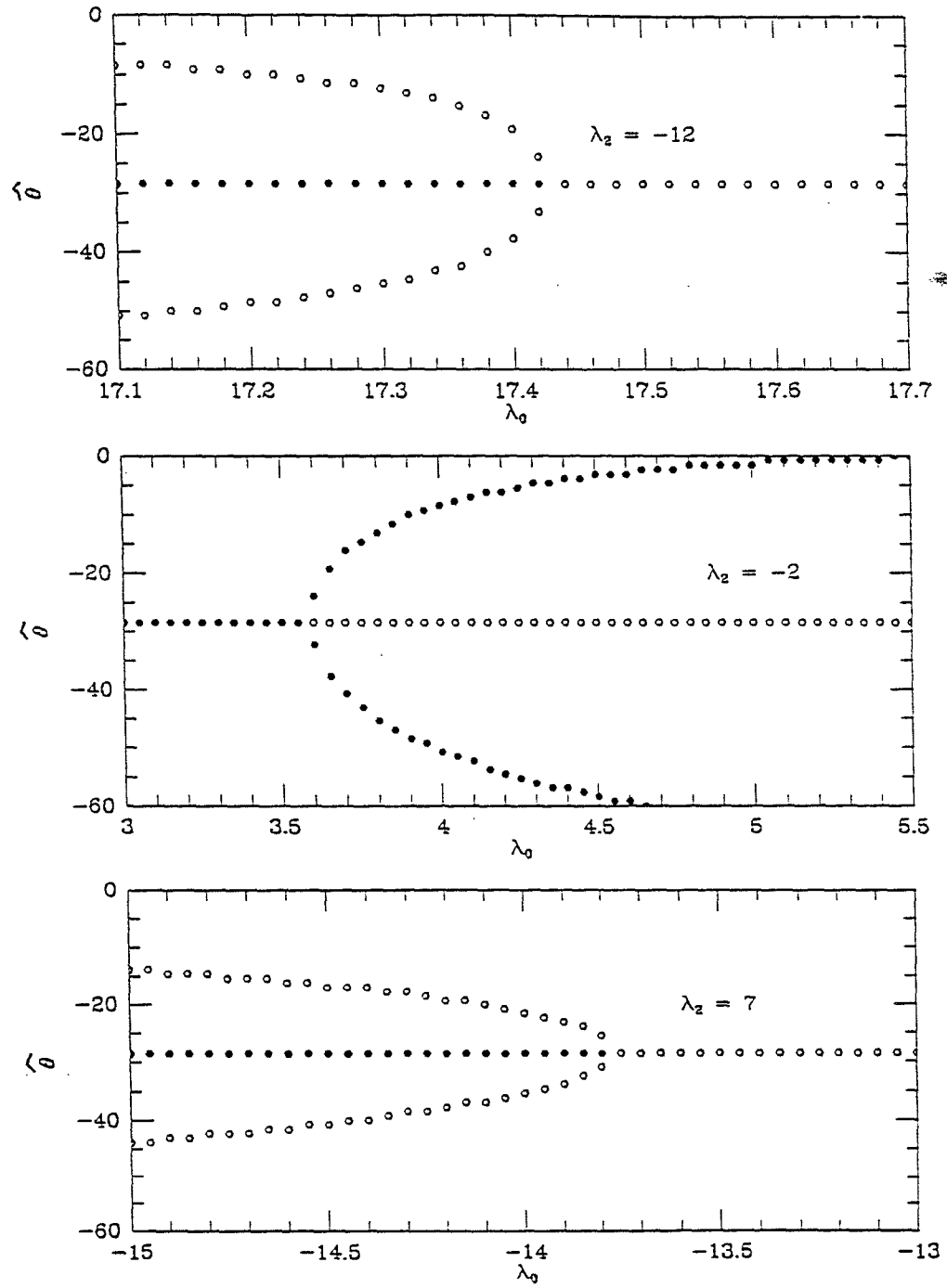
Figure 3.5.4 Local bifurcation diagrams

Appendix 3.A

In this appendix, we review some basic concepts and results of bifurcation theory. The standard reference is [14].

Consider a single scalar equation

$$g(x, \lambda) = 0. \tag{3.A.1}$$

The bifurcation theory studies how the solutions x of this equation change with the parameter λ; or, more precisely, what type of bifurcation occurs with parameter λ. Without loss of generality, one can assume $g(0,0) = 0$. Moreover, we assume $g : \Re \times \Re \to \Re$ is smooth. This is one of the standard local (static) bifurcation problems with one state variable, called *the recognition problem*. As with many bifurcation problems, this problem can be solved successfully through singularity theory, in which the related issue is called *finite determinacy*. Let $z = (x, \lambda)$. Near origin, the function g can be written as

$$g(z) = \sum_{|\alpha| < k+1} \frac{1}{\alpha!} (\frac{\partial}{\partial z})^\alpha g(0) z^\alpha + \sum_{|\alpha| = k+1} a_\alpha(z) z^\alpha, \tag{3.A.2}$$

for some smooth functions a_α defined in a neighborhood of the origin. Here we used the conventions with multi-indices:

$$|\alpha| = \alpha_1 + \alpha_2, \qquad \alpha! = (\alpha_1)!(\alpha_2)!,$$

$$z^\alpha = x^{\alpha_1} \lambda^{\alpha_2}, \qquad (\frac{\partial}{\partial z})^\alpha = (\frac{\partial}{\partial x})^{\alpha_1} (\frac{\partial}{\partial \lambda})^{\alpha_2}.$$

A key question is what terms in (3.A.2) can be ignored such that the values of coefficients of remaining terms can be used to determine the qualitative behavior of the original equation (3.A.1), for example, the variation in number of solutions. Singularity theory solves this problem by finding a suitable change of coordinates such that function g is equivalent to a standard model h, called *normal form*. A precise definition is given as follows (see [14]).

Definition 3.A.1: Two smooth mappings $g, h : \Re \times \Re \to \Re$ defined near the origin are *equivalent* if there exist a local diffeomorphism of $\Re^2$, $(x, \lambda) \mapsto (X(x, \lambda), \Lambda(\lambda))$ at the origin and a nonzero function $S(x, \lambda)$, such that

$$g(x, \lambda) = S(x, \lambda) h(X(x, \lambda), \Lambda(\lambda)), \tag{3.A.3}$$

63

where $X_x(0,0) > 0$ and $\Lambda'(0) > 0$. If $\Lambda = \lambda$, g and h are *strongly equivalent.* ■

From this definition we see that, since $S(x, \lambda)$ is nonzero, the solution of $g(x, \lambda) = 0$ and $h(X, \Lambda) = 0$ are the same in the sense of diffeomorphism. From this point of view, by means of singularity theory one can show why and what the high-order terms in (3.A.2) do not effect the qualitative behavior of equation $g(x, \lambda) = 0$. It should be noticed that although this method does not tell us how to derive an appropriate normal form h, for most physical problems, such as the one considered in this paper, it is not hard to pick up some of the candidates from a large number of known simple polynomials of x and λ, or the model of normal forms which have standard bifurcation diagrams. This is in essence the spirit of application of singularity theory to a physical problem.

Without considering detailed issues of singularity theory which are applicable to bifurcation problems, we directly give the following result which will be used in the next section. For details see [14], chapter 2. First we need the concept of germs.

Definition 3.A.2: Two smooth functions defined near the origin are *equivalent as germs* if there is some neighborhood of the origin on which they coincide. Let $\mathcal{E}_{x,\lambda}$ denote the set of equivalence classes of such functions. The elements in $\mathcal{E}_{x,\lambda}$ are called *germs.* ■

Lemma 3.A.3: A germ $g \in \mathcal{E}_{x,\lambda}$ is strongly equivalent to

$$\epsilon x^k + \delta \lambda x \qquad (3.A.4)$$

for $k > 2$ if and only if at $x = \lambda = 0$

$$g = \frac{\partial}{\partial x} g = \cdots = (\frac{\partial}{\partial x})^{k-1} g = \frac{\partial}{\partial \lambda} g = 0 \qquad (3.A.5a)$$

and

$$\epsilon = sgn(\frac{\partial}{\partial x})^k g, \qquad \delta = sgn \frac{\partial}{\partial x} \frac{\partial}{\partial \lambda} g. \qquad (3.A.5b)$$

■

Remark 3.A.4: When $k = 3$ the normal form (3.A.4) provides a pitchfork bifurcation. From this lemma, so is g if (3.A.5) holds. It is easy to show that if $\epsilon\delta > 0$, the pitchfork bifurcation is subcritical; if $\epsilon\delta < 0$, the pitchfork bifurcation is supercritical. ■

64

CHAPTER IV

OPTIMAL CONTROL PROBLEM ON
A RIGID BODY WITH TWO OSCILLATORS

An interesting problem in multibody mechanics is the problem of nonholonomic motion planning, or kinematic control problem. In recent research on various multibody mechanical systems with symmetry, the theory of principal bundles with connections has led to clear insight into the geometric structure of the problem, and provided a common framework for the formulation of related optimal control problem. However, explicit or partially explicit solution to the necessary conditions, given by differential equations on phase space, for the optimal path and control is still a challenge. Although, under certain conditions, the symmetries of the systems imply the existence of conserved quantities for the differential equations given by the necessary conditions, working with local coordinates at an early stage of the analysis usually causes difficulties in uncovering such quantities. In this chapter, we consider a particular mechanical system consisting of a rigid body and two point-masses, for which the structure group of the principal bundle is non-Abelian. For such mechanical systems, formulating the necessary condition for the optimal control in an intrinsic way is no longer trivial. We derive this condition by a

65

perturbation method and a Hamiltonian formulation and, based on structure symmetry and localization, explore the explicit solvability of optimal control for this system from a Hamiltonian point of view. The procedures of symplectic and Poisson reduction are applied systematically for this purpose.

4.1 Preliminaries

In this section, we give some mathematical background for this and next chapters. We first recall some useful definitions and results about principal fiber bundles and the theory of connections in differential geometry. The main references for this part are [6,33]. Then, we show an important principal connection for the simple mechanical system with symmetry, i.e., mechanical connection, which is originally due to Smale and Kummer (cf. [24]).

4.1.1 Principal Fiber Bundle and Connections

Definition 4.1.1: Let B and Q be smooth manifolds, referred to as the *base space* and the *total space*, respectively. Let G be a Lie group, referred to as *structure group*. A four-tuple $\wp = (Q, B, \pi, G)$ is called a *principal fiber bundle* or *principal G-bundle* if the following conditions are satisfied:

(a) G acts on Q to the left, freely and differentiably, where the action is denoted by

$$\Phi : G \times Q \to Q$$
$$(g, q) \mapsto \Phi(g, q) \stackrel{\triangle}{=} \Phi_g(q) \stackrel{\triangle}{=} g \cdot q; \tag{4.1.1}$$

(b) B is the quotient space of Q by the equivalence relation induced by G, i.e., $B = Q/G$, and the canonical projection $\pi : Q \to B$ is differentiable;

(c) Q is locally trivial, that is, for each $x \in B$ there is a neighborhood U of x such that $\pi^{-1}(U)$ is isomorphic with $U \times G$ in the sense that $q \in \pi^{-1}(U) \mapsto (\pi(q), \phi(q))$ is a diffeomorphism, where $\phi : \pi^{-1}(U) \to G$ satisfies $\phi(g \cdot q) = g\phi(q), \forall g \in G$. ∎

Remark 4.1.2: In the traditional definition of principal fiber bundle, the structure group G acts on Q to the *right*. In general, the choice of left or right actions is related to the properties of the physical system under study. Therefore in this and next chapters, we will also show the corresponding formulations for right action whenever it is necessary.

■

For $x \in B$, *the fiber over* x is a closed submanifold of Q which is differentiably isomorphic with G. For any point $q \in Q$, the *fiber through* q is the fiber over $x = \pi(q)$. The *trivial* or *product* principal fiber bundle is a special principal fiber bundle, in which $Q = B \times G$. In this case, G acts on Q by $\Phi(g, (x, h)) = (x, gh)$ for $x \in B$ and $g, h \in G$.

Let V be an r-dimensional vector space with basis $v_1, \cdots, v_r$. If $\alpha_1, \cdots, \alpha_r$ belong to $\varpi^k(Q)$, the space of real-valued k-forms, then $\sum_{i=1}^{r} \alpha_r v_r$ is called a V-*valued k-form*. The space of V-valued k-forms on Q is denoted by $\varpi^k(Q; V)$. Let $\omega \in \varpi^k(Q; V)$. As a real-valued k-form, its exterior derivative satisfies the following important equation (Cartan's formula):

$$
\begin{aligned}
d\omega(X_0, \cdots, X_k) = &\sum_{i=0}^{k} (-1)^i X_i(\omega(X_0, \cdots, \check{X}_i, \cdots, X_k)) \\
&+ \sum_{i<j} (-1)^{i+j} \omega([\, X_i, X_j\,], X_0, \cdots, \check{X}_i, \cdots, \check{X}_j, \cdots, X_k),
\end{aligned}
\tag{4.1.2}
$$

where $X_i \in \mathfrak{X}(Q), i = 0, \cdots, k$ and $\check{X}_i$ denotes that X_i is deleted. In particular, if $\omega \in \varpi^1(Q; V)$,

$$
(d\omega)(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([\, X, Y\,]),
\tag{4.1.3}
$$

where $X, Y \in \mathfrak{X}(Q)$. Note that, here we used the convention of exterior derivative in [1], instead of the one in [33].

We now ready to define connections on a principal fiber bundle.

Definition 4.1.3: A *connection* on a principal fiber bundle $\wp = (Q, B, \pi, G)$ is an assignment to each point $q \in Q$ a tangent subspace $\mathbf{H}_q \subset T_q Q$ such that, for $\mathbf{V}_q \triangleq \{v \in T_q Q \mid T\pi(v) = 0\}$, we have

(a) $T_q Q = \mathbf{H}_q \oplus \mathbf{V}_q$;

(b) for each $g \in G$ and $q \in Q$, $T_q \Phi_g \cdot \mathbf{H}_q = \mathbf{H}_{g \cdot q}$;

(c) $\mathbf{H}_q$ depends differentiably on q. ∎

In the above definition, the subspace $\mathbf{V}_q$ is called the vertical subspace of $T_q Q$ and $\mathbf{H}_q$ the horizontal subspace. Figure 4.1.1 gives a clear picture of the definition of connection.
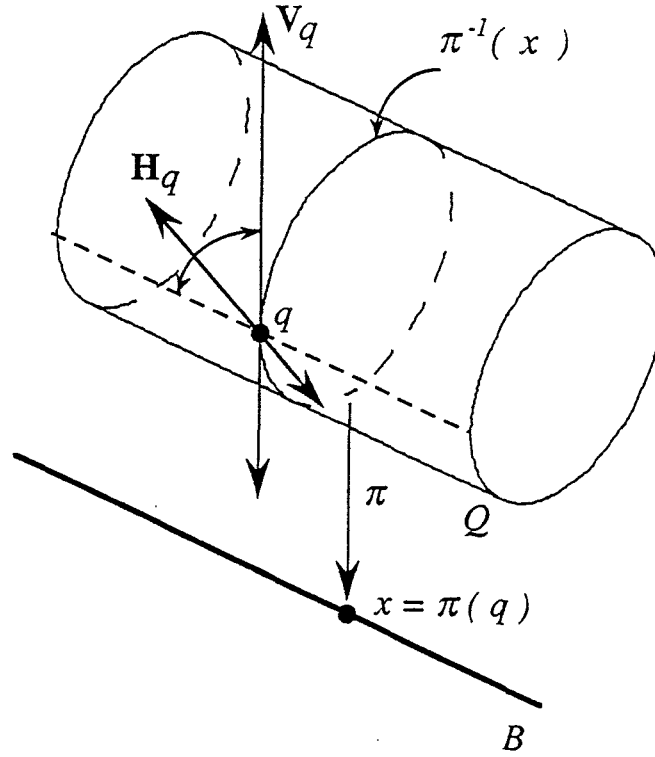
Figure 4.1.1 A Connection (cf. [6])

From the above definition of connection, the decomposition of the tangent space of Q leads to the decomposition of the vector fields on Q as follows. Let X be a smooth vector field on Q, from condition (a) in the above definition, at each point $q \in Q$,

$$X(q) = X^v(q) + X^h(q),$$

where $X^v(q) \in \mathbf{V}_q$ and $X^h(q) \in \mathbf{H}_q$. From the condition (c), by associating to each point q the tangent vector $X^v(q)$ and $X^h(q)$, we get smooth vector fields X^v and X^h which are called vertical and horizontal component of vector field X, respectively. If $X^v = X$ or $X^h = X$, X is called *vertical or horizontal vector field*, respectively.

From the properties of horizontal subspace, we see that a connection defines a subbundle over Q of TQ, i.e., *a (horizontal) distribution* on Q. Moreover, since the fiber over any point in Q is differentiably isomorphic with Lie group G, it is easy to verify that $\mathbf{V}_q = \{\xi_Q(q) | \xi \in \mathcal{G}\}$, where ξ_Q is infinitesimal generator of action given in (4.1.1) with respect to $\xi \in \mathcal{G}$ or *fundamental vector field on* Q, and the vertical vector

68

field X^v can be represented as, at each point $q \in Q$,

$$X^v(q) = [\xi(q)]_Q(q),$$

where $\xi(q)$ is a choice of an element in $\mathcal{G}$ at q.

An alternative way to define a connection is as follows.

Definition 4.1.4: Let $\mathcal{G}$ be the Lie algebra of G. A connection on a principal fiber bundle $\wp = (Q, B, \pi, G)$ is a $\mathcal{G}$-valued 1-form on Q, $\omega \in \varpi^1(Q; \mathcal{G})$, such that

(1) $\omega(\xi_Q(q)) = \xi$, for any $\xi \in \mathcal{G}$;

(2) $((\Phi_g)^*\omega)(X) = Ad_g\omega(X)$ for any $X \in \mathfrak{X}(Q)$.

A $\mathcal{G}$-valued 1-form ω satisfying (1) and (2) is referred to as the *connection form*. ∎

Remark 4.1.5: Because of the condition (2) in Definition 4.1.4, we say that the connection form is *Ad-equivariant*. If the Lie group G acts on Q to the *right*, the condition (2) becomes

$$((\Phi_g)^*\omega)(X) = Ad_{g^{-1}}\omega(X), \qquad \forall X \in \mathfrak{X}(Q). \qquad ∎$$

One can show that Definition 4.1.3 and 4.1.4 are equivalent (cf. [6]), that is, given a connection form, ω, there exists an unique decomposition of T_qQ or a choice of horizontal subspace $\mathbf{H}_q = Ker(\omega(q))$ at each point $q \in Q$; conversely, given a choice of horizontal subspace, there exists an unique $\mathcal{G}$-valued 1-form satisfies the conditions in Definition 4.1.4.

One should note that, for a given principal fiber bundle, there exists an infinity of choice of horizontal subspace/connection. This obvious claim turns out to be important in the next chapter where we consider a family of splittings of the dynamics of Lagrangian system.

Let $\omega \in \varpi^k(Q; \mathcal{G})$. The *covariant derivative* of ω, $\mathbf{D}\omega \in \varpi^{k+1}(Q; \mathcal{G})$, is defined by

$$\mathbf{D}\omega(X_0, \cdots, X_k) \overset{\triangle}{=} d\omega(X_0^h, \cdots, X_k^h),$$

where X_i^h is horizontal component of vector field $X_i \in \mathfrak{X}(Q)$. If $\omega \in \varpi^1(Q; \mathcal{G})$ is the connection form of a connection, the covariant derivative of ω, $\Omega = \mathbf{D}\omega$, is called

curvature form of the connection. We give three important properties of the curvature form below. For proof, see [33].

Proposition 4.1.6: Let $\omega \in \varpi^1(Q; \mathcal{G})$ be the connection form and Ω be the corresponding curvature form. Let X, Y be any vector fields on Q. Then, we have

(1) $((\Phi_g)^* \Omega)(X, Y) = Ad_g \Omega(X, Y)$;

(2) Structural equation:

$$\Omega(X, Y) = d\omega(X, Y) + [\,\omega(X), \omega(Y)\,]; \qquad (4.1.4)$$

(3) $\Omega(X, Y) = -\omega([\,X^h, Y^h\,])$. ∎

If a connection is determined on a principal fiber bundle $\wp = (Q, B, \pi, G)$, certain types of curves on Q can be characterized. A smooth curve $\{q(t), t \in [0, 1]\}$ on Q is called a *horizontal curve* if its tangent vectors are all horizontal, i.e., $\frac{dq(t)}{dt} \in \mathbf{H}_{q(t)}, \forall t \in [0, 1]$. Let $\{x(t), t \in [0, 1]\}$ be a piecewise smooth curve in B. A *horizontal lift* of $x(\cdot)$, $\{q(t), t \in [0, 1]\}$, is a curve in Q which is horizontal and $\pi(q(t))) = x(t), \forall t \in [0, 1]$. The following theorem is crucial for understanding the holonomy or the geometric phases.

Theorem 4.1.7: Let $\{x(t), t \in [0, 1]\}$ be a curve in B and let q_0 be any point in Q such that $\pi(q_0) = x(0)$. Then, there is a unique horizontal lift $\{q(t), t \in [0, 1]\}$ which starts at $q(0) = q_0$. ∎

The proof of this theorem uses the existence and uniqueness of the solution of an ordinary differential equation and can be found in [33].

From Theorem 4.1.7, one can define a mapping, called *parallel displacement*, $\psi : q_0 = q(0) \mapsto q_1 = q(1)$. It is easy to check that parallel displacement is a differential isomorphism of the fiber through $q(0)$ onto the fiber through $q(1)$ and it commutes with the group action Φ, i.e., $\psi(\Phi(g, q)) = \Phi(g, \psi(q))$ for any $g \in G$ and $q \in Q$.

Consider a closed curve in B, $\{x(t), t \in [0, 1]\}$ with $x(0) = x(1) \stackrel{\triangle}{=} x_0$. The parallel displacement now is an automorphism of the fiber over x_0, $\pi^{-1}(x_0)$. The set of all of such automorphisms corresponding to all closed curves at x_0 in B forms a group, called the *holonomy group* at x_0, denoted by Ψ_{x_0}. One shows that Ψ_{x_0} can be identified as a subgroup of the structure group G. An element in Ψ_{x_0} is called a *holonomy* at x_0.

Given a connection on a principal fiber bundle $\wp = (Q, B, \pi, G)$, we now can

compute the holonomy corresponding to a closed curve in B, say, $x(\cdot) = \{x(t), t \in [0,1]\}$, and a point $q_0 \in \pi^{-1}(x(0)) \subset Q$. Assume $\{x(t), t \in [0,1]\}$ is contained in an open set U of B. Let $\sigma : U \to Q$ be a local section of the bundle and $\omega \in \varpi^1(Q; \mathcal{G})$ be the connection form. Then the holonomy at q_0 with respect to $x(\cdot)$ is the solution of the following differential equation on G at $t = 1$:

$$\frac{dg(t)}{dt} = -T_e L_g \cdot (\sigma^* \omega)(\dot{x}(t)). \tag{4.1.5}$$

Indeed, let $q(\cdot) = \{q(t), t \in [0,1]\}$ be the horizontal lift of $x(\cdot)$ with $q(0) = q_0$ and $g(\cdot) = \{g(t), t \in [0,1]\}$ be a curve in G such that $q(t) = \Phi(g(t), \sigma(x(t))) = \Phi_{g(t)}(\sigma(x(t))), \forall t \in [0,1]$. Then,

$$\frac{dq(t)}{dt} = T_{\sigma(x(t))} \Phi_{g(t)} T_{x(t)} \sigma \dot{x}(t) + T_{\sigma(x(t))} \Phi_{g(t)} \xi_Q(\sigma(x(t))),$$

where $\xi \stackrel{\triangle}{=} T_g L_{g^{-1}} \dot{g}(t) \stackrel{\triangle}{=} g(t)^{-1} \cdot \dot{g}(t)$. Let us apply the connection form to both sides of the above equation. Since $q(\cdot)$ is horizontal and ω is Ad-equivariant, we have

$$0 = (\sigma^* \omega)(\dot{x}(t)) + \xi. \tag{4.1.6}$$

From the definition of ξ, we get Equation (4.1.5).

The $\mathcal{G}$-valued form, $\sigma^* \omega \in \varpi^1(B, \mathcal{G})$, is called the *local connection form*. From (4.1.5), if G is an Abelian group, the holonomy, or $g(1)$, can be represented explicitly as

$$g(1) = exp(-\int_0^1 (\sigma^* \omega)(\dot{x}(t)) dt) = exp(-\iint_{\mathcal{D}} \sigma^* \Omega), \tag{4.1.7}$$

where $\mathcal{D}$ is a surface in B with $x(\cdot)$ as the boundary and Ω is the curvature form of the connection.

In next subsection, we consider an application of the theory introduced here to simple mechanical systems with symmetry.

4.1.2 Mechanical Connection and Related Control Problems

With the background of differential geometry introduced in the preceding subsection, we now ready to give the basic ingredients in formulating the problems of geometric

phases and optimal control in mechanics. Consider a simple mechanical system with symmetry (cf. Subsection 2.1.4), (Q, K, V, G), together with an equivariant momentum map $\mathbf{J} : TQ \to \mathcal{G}^*$ satisfying

$$\langle \mathbf{J}(q, v_q), \ \xi \rangle = K(q)(v_q, \xi_Q(q)) \qquad \forall \xi \in \mathcal{G}. \tag{4.1.8}$$

In addition, we also let Q be the total space of a principal G-bundle $\wp = (Q, B, \pi, G)$. Here, the base space B is also referred to as the *shape space*. On this bundle, the mechanical connection is constructed as follows. At each point $q \in Q$, define the *locked inertia tensor* as the mapping

$$\mathbb{I}(q) : \mathcal{G} \to \mathcal{G}^* \tag{4.1.9a}$$

such that

$$\langle \mathbb{I}(q)\eta, \ \xi \rangle = K(q)(\eta_Q(q), \xi_Q(q)) \quad \forall \eta, \xi \in \mathcal{G}. \tag{4.1.9b}$$

From this definition, it is obvious that $\langle \mathbb{I}(q)\eta, \ \xi \rangle = \langle \mathbb{I}(q)\xi, \ \eta \rangle$. This map is called as the locked inertia tensor since, for coupled rigid or elastic systems, it is the moment of inertia tensor of the relatively locked/frozen system (cf. Remark 3.5.5 (2)).

Defining a $\mathcal{G}$-valued 1-form $\alpha \in \varpi^1(Q; \mathcal{G})$ by

$$\alpha : TQ \to \mathcal{G}$$
$$(q, v) \mapsto \alpha(q, v) = \mathbb{I}^{-1}(q)(\mathbf{J}(q, v)), \tag{4.1.10}$$

one can show the following theorem due to Kummer and Smale.

Theorem 4.1.8: The $\mathcal{G}$-valued one-form α in (4.1.10) defines a connection, referred as *mechanical connection*, on the principal bundle $\wp = (Q, B, \pi, G)$.

Proof: We need to check conditions (1) and (2) in Definition 4.1.4. For any $\nu \in \mathcal{G}^*$,

$$\langle \nu, \ \alpha(\xi_Q(q)) \rangle = \langle \nu, \ \mathbb{I}^{-1}(q)(\mathbf{J}(\xi_Q(q))) \rangle$$
$$= \langle \mathbf{J}(\xi_Q(q)), \ \mathbb{I}^{-1}(q)\nu \rangle$$
$$= K(q)(\xi_Q, (\mathbb{I}^{-1}(q)\nu)_Q)$$
$$= \langle \mathbb{I}(q)(\mathbb{I}^{-1}(q)\nu), \ \xi \rangle = \langle \nu, \ \xi \rangle.$$

So, condition (1) is proved. To prove (2), we first show that the locked inertia tensor has the equivariance property:

$$\mathbb{I}(\Phi_g(q)) \cdot Ad_g\xi = Ad_{g^{-1}} \mathbb{I}(q) \cdot \xi. \tag{4.1.11}$$

For any $\eta \in \mathcal{G}$,

$$\langle \mathbb{I}(\Phi_g(q))Ad_g\xi, \;\; \eta \rangle = K(\Phi_g(q))((Ad_g\xi)_Q(\Phi_g(q)), \eta_Q(\Phi_g(q)))$$

$$= K(q)(T_{g \cdot q}\Phi_g^{-1}(Ad_g\xi)_Q(\Phi_g(q)), T_{g \cdot q}\Phi_g^{-1}\eta_Q(\Phi_g(q)))$$

$$= K(q)(\xi_Q(q), (Ad_{g^{-1}}\eta)_Q(q))$$

$$= \langle \mathbb{I}(q)\xi, \;\; Ad_{g^{-1}}\eta \rangle$$

$$= \langle Ad_{g^{-1}}^* \mathbb{I}(q)\xi, \;\; \eta \rangle.$$

So, (4.1.11) is proved. Here we used the relation (cf. Proposition 4.1.26 in [1])

$$T_{g \cdot q}\Phi_g^{-1}(\eta)_Q(\Phi_g(q)) = (Ad_{g^{-1}}\eta)_Q(q).$$

Now, let $\zeta \overset{\triangle}{=} (L_g^*\alpha)(q, v)$. Then,

$$\zeta = \alpha(\Phi_g(q), T_q\Phi_g(v)) = \mathbb{I}(g \cdot q)^{-1}\mathbf{J}(g \cdot q, T_q\Phi_g v) = \mathbb{I}(g \cdot q)^{-1}Ad_{g^{-1}}^*\mathbf{J}(q, v).$$

This implies $\mathbb{I}(g \cdot q)\zeta = Ad_{g^{-1}}^*\mathbf{J}(q, v)$. But, by (4.1.11),

$$\mathbb{I}(g \cdot q)\zeta = \mathbb{I}(g \cdot q)Ad_g(Ad_{g^{-1}}\zeta) = Ad_{g^{-1}}^* \mathbb{I}(Ad_{g^{-1}}\zeta).$$

Therefore, we have $\mathbf{J}(q, v) = \mathbb{I}(q)(Ad_{g^{-1}}\zeta)$ or $\zeta = Ad_g \mathbb{I}(q)^{-1}\mathbf{J}(q, v)$. Condition (2) is proved. ∎

From the connection theory introduced in the preceding subsection, when the connection form is defined, we have a vertical-horizontal splitting of the tangent bundle TQ. For the mechanical connection given in (4.1.10), we have

$$T_qQ = (Vert)_q \oplus (Hor)_q \tag{4.1.12a}$$

such that, for each $v_q \in T_qQ$,

$$v_q = (\alpha(v_q))_Q(q) + (v_q - (\alpha(v_q))_Q(q))$$
$$= (\mathbb{I}(q)^{-1}\mu)_Q(q) + (v_q - (\mathbb{I}(q)^{-1}\mu)_Q(q)), \tag{4.1.12b}$$

where $\mu = \mathbf{J}(v_q)$. It is readily shown that

$$Hor = \{(q, v) \in TQ | \mathbf{J}(q, v) = 0\}, \tag{4.1.13}$$

and the splitting in (4.1.12) is orthogonal one with respect to metric K.

To formulate the kinematic control problem explicitly, we consider the trivial bundle, i.e., $\wp = (B \times G, B, \pi, G)$. Here, the control is *internal* to the system, which leaves invariant the conserved momentum map **J**. Since, by definition, a principal fiber bundle is locally trivial, the equations we have below are locally true in general.

The tangent space at each point $(x, g) \in Q$ is represented by

$$T_{(x,g)}Q = T_x B \times T_g G$$

and let a tangent vector in $T_{(x,g)}Q$ be represented by $v_{(x,g)} = (v_x, v_g)_{(x,g)} = (v_x, g \cdot \xi)_{(x,g)}$, where $\xi = T_g L_{g^{-1}} v_g \in \mathcal{G}$. The Lie group G acts on Q following the rule $\Phi(h, (x, g)) = (x, hg)$, where $h, g \in G$ and $x \in B$. Then the infinitesimal generator corresponding to $\eta \in \mathcal{G}$ is

$$\eta_Q(q) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \Phi(exp(\epsilon\eta), (x, g)) = (0, \eta \cdot g). \tag{4.1.14}$$

Using the G-invariance of K, we have

$$\mathbf{J}(q, v) \cdot \eta = K(x, g)((v_x, g \cdot \xi), (0, \eta \cdot g))$$

$$= K(x, e)((v_x, \xi), (0, Ad_{g^{-1}}\eta))$$

$$= \langle \widetilde{\mathbb{I}}(x)\xi, \ Ad_{g^{-1}}\eta \rangle + \langle \jmath(x)(v_x), \ Ad_{g^{-1}}\eta \rangle$$

$$= \langle Ad^*_{g^{-1}}(\widetilde{\mathbb{I}}(x)\xi + \jmath(x)(v_x)), \ \eta \rangle,$$

where e is the identity element in G; $\widetilde{\mathbb{I}}(x) \stackrel{\triangle}{=} \mathbb{I}(x, e)$ is referred to as (local) locked inertia tensor at x, represents the metric on G, and $\jmath(x) : T_x B \to \mathcal{G}^*$ comes from the cross term when the metric K is written in terms of metrics on B and G. Here, the metric on B is induced from K. Therefore, we have, for $\mu = \mathbf{J}(q, v)$,

$$\mu = Ad^*_{g^{-1}}\widetilde{\mathbb{I}}(x)\xi + Ad^*_{g^{-1}}\jmath(x)(v_x), \tag{4.1.15}$$

or

$$\xi = \widetilde{\mathbb{I}}(x)^{-1} Ad^*_g \mu - \widetilde{\mathbb{I}}(x)^{-1} \jmath(x)v_x,$$

or, by left action,

$$v_g = g \cdot (\widetilde{\mathbb{I}}(x)^{-1} Ad^*_g \mu - \widetilde{\mathbb{I}}(x)^{-1} \jmath(x)v_x). \tag{4.1.16}$$

Given a closed curve $x(\cdot) = \{x(t), t \in [0, 1]\}$ in B and an initial point $q_0 = (x_0, g_0)$ in Q, using (4.1.16) one will be able to compute the shift, or the *phase*, in G. The phase

generated by the first term in (4.1.16) is referred to as a *dynamic phase* and the phase generated by the second term in (4.1.16) is the holonomy, referred to as *geometric phase*. One can show that $\widetilde{\mathbb{I}}(x)^{-1}\jmath(x)(\dot{x})$ is, in fact, the value of the local connection form of the mechanical connection at $(x, \dot{x})$.

Assuming that the vector $\dot{x}$ or the velocity of the path in B can be directly controlled, from (4.1.16), an associated kinematic control system can be set up as

$$\begin{cases} \dot{x} = u \\ \dot{g} = g \cdot (\widetilde{\mathbb{I}}(x)^{-1} Ad_g^* \mu - \widetilde{\mathbb{I}}(x)^{-1} \jmath(x) u), \end{cases} \qquad (4.1.17a)$$

or simply

$$\dot{q} = X_\mu(q) + \mathcal{H}(q)u, \qquad (4.1.17b)$$

for $q = (g, x) \in Q$, where $X_\mu(q) = (0, g \cdot \widetilde{\mathbb{I}}(x)^{-1} Ad_g^* \mu)$ is the drift, $\mathcal{H}(q) : T_{\pi(q)}(B) \rightarrow T_q Q$ is the horizontal lift operator and $u \in T_{\pi(q)}(B)$ is a tangent vector on shape space representing controls. Two control problems for this system can be framed as follows:

(P1) Given two points q_0 and q_1 in Q, find $u(\cdot)$ steering q_0 to q_1 at a specified time;

(P2) Given two points q_0 and q_1 in Q on the same fiber, find $u(\cdot)$ steering q_0 to q_1 while minimizing

$$\int_0^T < u, u >_B dt$$

for Riemannian metric $< \cdot, \cdot >_B$ on B and the fixed final time $T > 0$ subject to (4.1.17).

Remark 4.1.9: One should note that the vector field given in (4.1.17) is G_μ-invariant on the left, and not invariant under the action of the structure group G. In addition, It is obvious that for any control law, the solution of (4.1.17) satisfies $\mathbf{J}((x, g), (\dot{x}, \dot{g})) = \mu$. This means that one can equivalently construct the above control problems on principal fiber bundle $(Q' = B \times G_\mu, B, \pi', G_\mu)$, where $\pi' = \pi|_{Q'}$. On this bundle, the control system (4.1.17) becomes

$$\begin{cases} \dot{x} = u \\ \dot{g} = g \cdot (\mathbb{I}(x)^{-1} \mu - \mathbb{I}(x)^{-1} \jmath(x) u), \end{cases} \qquad (4.1.18)$$

which has symmetry under the action of structure group G_μ. ∎

The problems (P1) and (P2) are standard problems in control theory, namely, controllability and optimal control. (P1) is equivalent to the problem of accessibility and the corresponding condition is known via Chow's theorem [9]. In addition, if $\mu = 0$ and the system is accessible, (P2) is the isoholonomic problem in [31], or a special case of the problem of singular Riemannian/sub-Riemannian/nonholonomic geodesics [8,48].

In the next section, we will formulate the control system and corresponding optimal control problem for the system of a rigid body with two oscillators following the above procedure.

4.2 Momentum Map and Connection

In this section, after a complete derivation of the momentum map by means of standard method, we give an explicit expression of the mechanical connection for the system consisting of a rigid body with two oscillators.
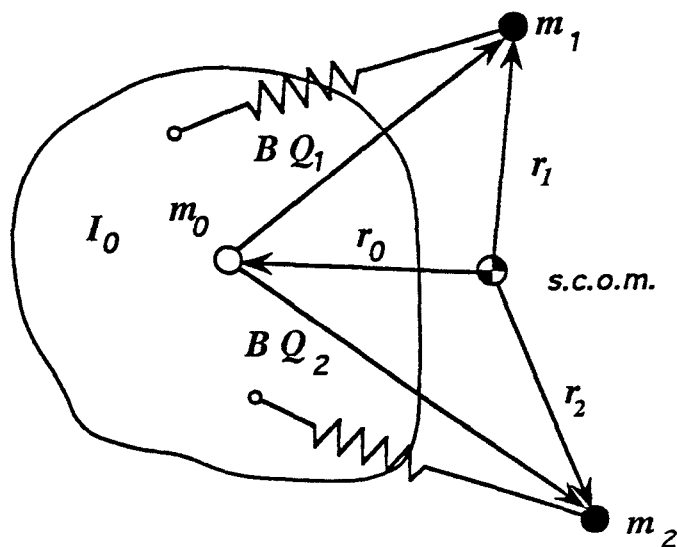
Figure 4.2.1 A Rigid body with two oscillators

The mechanical system we consider is shown in Figure 4.2.1. Here, r_0 is the position vector of the center of mass of the rigid body or *carrier* relative to the center of mass of the system; r_1 and r_2 are the position vectors of two oscillators with point masses m_1 and m_2 relative to the center of mass of the system, respectively; the mass and moment of inertia tensor of the carrier are denoted as m_0 and I_0, respectively; Q_1 and Q_2 are the position vectors of two oscillators relative to a frame (not displayed) fixed on carrier and $A \in SO(3)$ determines the orientation of the carrier with respect to an inertial frame (not displayed). We assume that no exterior force/torque affects the system and the potential energy is zero. This implies that the inertial frame can be placed at the center of mass of the system and r_0, r_1 and r_2 are related by

$$\sum_{i=0}^{2} m_i r_i = 0. \tag{4.2.1}$$

For now, r_1 and r_2 (or Q_1 and Q_2) are assumed to be arbitrarily time dependent vectors. Later, we will impose constraints on them to study the effect of their motion on the motion of the carrier.

From the above setting, we have the configuration space $Q = (\Re^3)^2 \times SO(3)$ with local coordinates $q = (r_1, r_2, A)$ and its tangent bundle $TQ = (T\Re^3)^2 \times TSO(3)$ with local coordinates $(q, v) = ((r_1, r_2, A), (\dot{r}_1, \dot{r}_2, A\widehat{\Omega}))$. Here, denoting by Ω the vector of angular velocity of the carrier with respect to the body fixed frame, we used the fact

$$\dot{A} = A\widehat{\Omega} \tag{4.2.2}$$

with the standard isomorphism

$$\begin{aligned}
\widehat{} : \Re^3 &\to so(3) \\
(x_1, x_2, x_3) &\mapsto \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}.
\end{aligned} \tag{4.2.3}$$

The Lagrangian of the system can be determined easily as

$$\begin{aligned}
L((r_1, r_2, A), (\dot{r}_1, \dot{r}_2, A\widehat{\Omega})) = &\frac{1}{2} < \Omega, I_0\Omega > + \frac{1}{2}(m_1 + \frac{m_1^2}{m_0}) < \dot{r}_1, \dot{r}_1 > \\
&+ \frac{1}{2}(m_2 + \frac{m_2^2}{m_0}) < \dot{r}_2, \dot{r}_2 > + \frac{m_1 m_2}{m_0} < \dot{r}_1, \dot{r}_2 >,
\end{aligned} \tag{4.2.4}$$

where $< \cdot, \cdot >$ is the inner product in $\Re^3$. This Lagrangian is given by a Riemannian metric K on Q, i.e., $L(q,v) = \frac{1}{2}K(q)(v,v)$, where, for $(u_1, u_2, A\widehat{\Omega})$ and $(w_1, w_2, A\widehat{\Xi}) \in T_{(r_1,r_2,A)}Q$,

$$K(r_1, r_2, A)((u_1, u_2, A\widehat{\Omega}), (w_1, w_2, A\widehat{\Xi})) \triangleq$$

$$< \Omega, I_0\Xi > +(m_1 + \frac{m_1^2}{m_0}) < u_1, w_1 > +(m_2 + \frac{m_2^2}{m_0}) < u_2, w_2 > \qquad (4.2.5)$$

$$+ \frac{m_1 m_2}{m_0} < u_1, w_2 > + \frac{m_1 m_2}{m_0} < u_2, w_1 > .$$

Let $G = SO(3)$ act on Q by

$$\Phi : SO(3) \times ((\Re^3)^2 \times SO(3)) \to (\Re^3)^2 \times SO(3)$$

$$(A, (r_2, r_2, B)) \mapsto (Ar_1, Ar_2, AB). \qquad (4.2.6)$$

From (4.2.5), one can show that G acts on Q by isometries. Therefore, by definition (cf. Subsection 2.1.4), the system $(Q = (\Re^3)^2 \times SO(3), K, V = 0, G = SO(3))$ is a simple mechanical system with symmetry.

By direct or some intrinsic calculations on $SO(3)$ (cf. [49]), one finds the Legendre transform, at $(q,v) \in TQ$, as

$$K^\flat(q)(v) = D_2 L(q,v) = (A\widehat{I_0\Omega}, (m_1 + \frac{m_1^2}{m_0})\dot{r}_1 + \frac{m_1 m_2}{m_0}\dot{r}_2, (m_2 + \frac{m_2^2}{m_0})\dot{r}_2 + \frac{m_1 m_2}{m_0}\dot{r}_1).$$

The infinitesimal generator of the action in (4.2.6) corresponding to $\widehat{\xi} \in so(3)$ is

$$\xi_Q(q) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \Phi(e^{\epsilon\widehat{\xi}}, (r_1, r_2, A)) = (\widehat{\xi}r_1, \widehat{\xi}r_2, \widehat{\xi}A).$$

Then from (4.1.8), the momentum map is calculated as follows:

$$\langle \mathbf{J}(q,v), \ \xi \rangle = \langle K^\flat(q)(v), \ \xi_Q \rangle$$

$$= \langle A\widehat{I_0\Omega}, \ \widehat{\xi}A \rangle + \langle (m_1 + \frac{m_1^2}{m_0})\dot{r}_1 + \frac{m_1 m_2}{m_0}\dot{r}_2, \ \widehat{\xi}r_1 \rangle$$

$$+ \langle (m_2 + \frac{m_2^2}{m_0})\dot{r}_2 + \frac{m_1 m_2}{m_0}\dot{r}_1, \ \widehat{\xi}r_2 \rangle$$

$$= \langle AI_0\Omega + (m_1 + \frac{m_1^2}{m_0})r_1 \times \dot{r}_1 + \frac{m_1 m_2}{m_0}r_1 \times \dot{r}_2$$

$$+ (m_2 + \frac{m_2^2}{m_0})r_2 \times \dot{r}_2 + \frac{m_1 m_2}{m_0}r_2 \times \dot{r}_1, \xi \rangle. \qquad (4.2.7)$$

For vector space $T_B SO(3)$ for some $B \in SO(3)$, pairing $\langle \cdot, \cdot \rangle$ is defined as

$$\langle \ , \ \rangle : T^* SO(3) \times TSO(3) \to \Re$$

$$(A_1, A_2) \mapsto \frac{1}{2}trace(A_1 A_2^T). \qquad (4.2.8)$$

In obtaining (4.2.7), we also used the facts:

$$\widehat{x}y = x \times y, \qquad \forall x, y \in \Re^3,$$

and

$$A\widehat{x}A^T = \widehat{Ax}, \qquad \forall x \in \Re^3, \forall A \in SO(3).$$

Letting $\mu = \mathbf{J}(q, v)$, from (4.2.7) we have

$$\mu = AI_0\Omega + (m_1 + \frac{m_1^2}{m_0})r_1 \times \dot{r}_1 + \frac{m_1 m_2}{m_0}r_1 \times \dot{r}_2$$
$$+ (m_2 + \frac{m_2^2}{m_0})r_2 \times \dot{r}_2 + \frac{m_1 m_2}{m_0}r_2 \times \dot{r}_1. \tag{4.2.9}$$

One can show that μ is, in fact, the total angular momentum of the system.

It is clear that $((\Re^3)^2 \times SO(3))$ is a trivial bundle with the structure group $SO(3)$ and the base space $(\Re^3)^2$ coordinatized by (Q_1, Q_2). Using coordinates (Q_1, Q_2, A) for the configuration space Q, the angular momentum in (4.2.9) can be rewritten as follows. From Figure 4.2.1 and Equation (4.2.1), we have

$$r_i = r + AQ_i, \qquad i = 1, 2$$

and

$$r = -A(\epsilon_1 Q_1 + \epsilon_2 Q_2),$$

where $\epsilon_i = \frac{m_i}{m_0 + m_1 + m_2}, i = 1, 2$. Equation (4.2.9) can be rearranged as

$$\mu = A((I_0 + \Delta I_0)\Omega + D_1\dot{Q}_1 + D_2\dot{Q}_2), \tag{4.2.10}$$

where

$$\Delta I_0 = -m(\epsilon_1 \widehat{Q}_1^2 + \epsilon_2 \widehat{Q}_2^2 - (\epsilon_1 \widehat{Q}_1 + \epsilon_2 \widehat{Q}_2)^2)$$

$$D_1 = m[(1 - \epsilon_1)\epsilon_1 \widehat{Q}_1 - \epsilon_1 \epsilon_2 \widehat{Q}_2]$$

$$D_2 = m[-\epsilon_1 \epsilon_2 \widehat{Q}_1 + \epsilon_2(1 - \epsilon_2)\widehat{Q}_2].$$

By (4.2.10), we have

$$\Omega = (I_0 + \Delta I_0)^{-1}(A^T\mu - (D_1\dot{Q}_1 + D_2\dot{Q}_2)) \tag{4.2.11}$$

or, by (4.2.2),

$$\dot{A} = A[(I_0 + \Delta I_0)^{-1}A^T\mu - (I_0 + \Delta I_0)^{-1}\widehat{(D_1\dot{Q}_1 + D_2\dot{Q}_2)}], \tag{4.2.12}$$

79

where $[\cdot]\widehat{} \triangleq \widehat{(\cdot)}$. Comparing (4.2.12) with (4.1.16), we see that

$$I_{lock}(Q_1,Q_2) \triangleq I_0 + \Delta I_0$$

is the (local) locked inertia, and

$$\omega(Q_1,Q_2)(\dot{Q}_1,\dot{Q}_2) \triangleq [(I_0 + \Delta I_0)^{-1}(D_1\dot{Q}_1 + D_2\dot{Q}_2)]\widehat{}$$

is the value of the local connection form at point $((Q_1,Q_2),(\dot{Q}_1,\dot{Q}_2)) \in TB$ with respect to the mechanical connection. This connection form can be explicitly given by

$$\begin{aligned}\breve{\omega}(Q_1,Q_2) &= I_{lock}^{-1}(D_1 dQ_1 + D_2 dQ_2)\\ &= mI_{lock}^{-1}[((1-\epsilon_1)\epsilon_1\widehat{Q}_1 - \epsilon_1\epsilon_2\widehat{Q}_2)dQ_1 \quad\quad (4.2.13)\\ &\quad\quad + (-\epsilon_1\epsilon_2\widehat{Q}_1 + \epsilon_2(1-\epsilon_2)\widehat{Q}_2)dQ_2],\end{aligned}$$

where the operator $\breve{}$ is the inverse of operator $\widehat{}$ (cf. (4.2.3)).

Equation (4.2.12) can be used for computing the phases of the system and the related optimal control problem mentioned in the preceding section. In particular, when $\mu = 0$, it can be used to compute holonomy, or geometric phase and to solve the isoholonomy problem. It should be noted that, in this case, the angular velocity vector of the rigid body in a body fixed frame is related to the connection form by

$$\begin{aligned}\Omega &= -\breve{\omega}(Q_1,Q_2)(\dot{Q}_1,\dot{Q}_2)\\ &= mI_{lock}^{-1}[((1-\epsilon_1)\epsilon_1\widehat{Q}_1 - \epsilon_1\epsilon_2\widehat{Q}_2)\dot{Q}_1 \quad\quad (4.2.14)\\ &\quad\quad + (-\epsilon_1\epsilon_2\widehat{Q}_1 + \epsilon_2(1-\epsilon_2)\widehat{Q}_2)\dot{Q}_2].\end{aligned}$$

In the following sections, we will consider the case $\mu = 0$ only. In addition, we will assume that the oscillators are confined to move along certain guide-ways. Under this assumption, the bundle structure will be simplified and the equations for phases and the connection form on such a bundle can be easily derived from those we have found.

4.3 Planar System

We now assume that the vector r, BQ_1 and BQ_2 are kept in the same plane in inertial space and that $m_1 = m_2$ (so $\epsilon_1 = \epsilon_2 \triangleq \epsilon$). In addition, we choose a body-fixed coordinate system (or frame) $0\text{-}xyz$ on the carrier with $0\text{-}z$ axis perpendicular to the

plane and the origin of this frame is placed at the center of mass of the carrier (see Figure 4.3.1). We also let the two oscillators move along two parallel guide-ways such that, in the 0-xyz frame,

$$Q_1 = (-l, x_1(t), 0)^T \quad \text{and} \quad Q_2 = (\, l, x_2(t), 0)^T,$$

where $x_1(t), x_2(t) \in \Re$ and l is the distance of the guide-ways to the origin of 0-xyz frame. It is clear that the configuration space is reduced to $Q = \Re^2 \times S^1$, which will be coordinatized by (x_1, x_2, θ), and the principal fiber bundle is $(\Re^2 \times S^1, \Re^2, \pi, S^1)$.
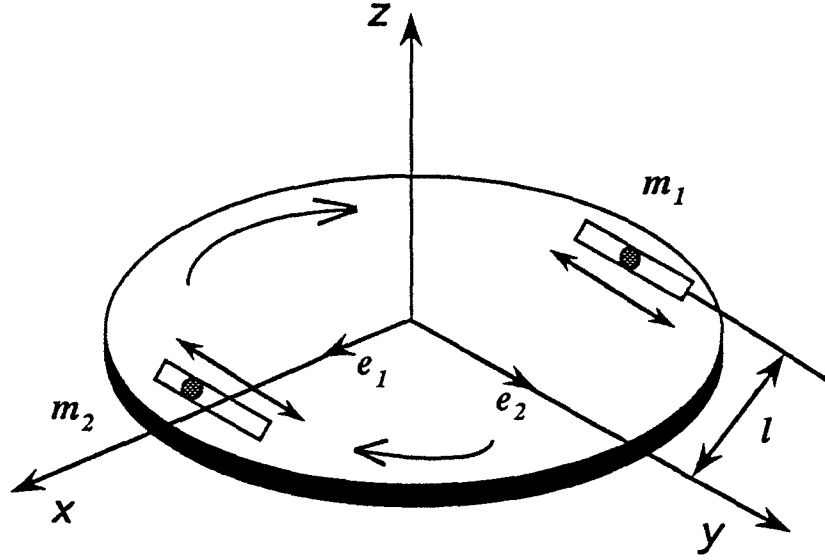
Figure 4.3.1 A rigid body with two oscillators: planar case

Setting $\mu = 0$, the angular velocity $\Omega = (\Omega_x, \Omega_y, \Omega_z)^T$ in (4.2.14) is of the form

$$\begin{cases} \Omega_x = 0 \\[2mm] \Omega_y = 0 \\[2mm] \Omega_z = \dot\theta = \dfrac{m\epsilon l}{I_{lock}}(\dot x_1 - \dot x_2) \end{cases} \qquad (4.3.1)$$

where

$$I_{lock} \triangleq I_z + m\epsilon(2l^2 + (1-\epsilon)x_1^2 - 2\epsilon x_1 x_2 + (1-\epsilon)x_2^2)$$

81

with I_z the moment of inertia of the carrier about z axis. It is obvious that the local connection form corresponding to the mechanical connection on the principal bundle $(\Re^2 \times S^1, \Re^2, \pi, S^1)$ is

$$\omega(x_1, x_2) = -\frac{m\epsilon l}{I_{lock}}(dx_1 - dx_2). \qquad (4.3.2)$$

For simplicity, we further assume that the amplitude of motion of each oscillator is very small in comparison with the the spacing of two guide-ways, i.e.

$$|x_i|/l << 1. \qquad (4.3.3)$$

Under this assumption, using Taylor expansion (up to quadratic terms of x_i), we get an approximate ω (with the same notation)

$$\omega = -\frac{m\epsilon l(I_z + 2m\epsilon l^2 - m\epsilon(2l^2 + (1-\epsilon)x_1^2 - 2\epsilon x_1 x_2 + (1-\epsilon)x_2^2))}{(I_z + 2m\epsilon l^2)^2}(dx_1 - dx_2).$$

The above procedure is called *localization* in [23]. Since we are interested in the motion of the carrier generated by the motion of the point masses on a *closed curve* in shape space, the above ω can be further simplified as follows. Applying exterior derivative to the above equation, we have

$$\begin{aligned}
d\omega &= -\frac{2m\epsilon l}{(I_z + 2m\epsilon l^2)^2}(x_1 + x_2)dx_1 \wedge dx_2 \\
&= -\frac{m\epsilon l}{(I_z + 2m\epsilon l^2)^2}d(x_1^2 dx_2 - x_2^2 dx_1).
\end{aligned}$$

Then, under the assumption (4.3.3), we can take (for closed paths in shape space)

$$\omega = -\frac{m\epsilon l}{(I_z + 2m\epsilon l^2)^2}(x_1^2 dx_2 - x_2^2 dx_1), \qquad (4.3.4)$$

modulo an exact one-form.

Let $c(\cdot)$ be any closed curve in shape space $\Re^2$. Since S^1 is Abelian, from (4.1.7), the corresponding geometric phase or holonomy, i.e. the drift rate of the carrier about z-axis, will be

$$\delta_z = -\int_c \omega = -\int_c \frac{m\epsilon l}{(I_z + 2m\epsilon l^2)^2}(x_1^2 dx_2 - x_2^2 dx_1). \qquad (4.3.5)$$

Using (4.3.5), we now compute the geometric phase for the case in which both oscillators follow sinusoidal motions with different amplitudes, frequencies and phase angles.

Let

$$x_1(t) = a_1 sin(\bar{\omega}t + \phi_1)$$

$$x_2(t) = a_2 sin(n\bar{\omega}t + \phi_2)$$

for $t \in [0, 2\pi/\bar{\omega}]$, where $\bar{\omega}, a_i, \phi_i$ are real numbers and n is an integer. Then, the closed curve in the shape space forms a Lissajous figure. Substituting the above $x_i(t)$ into (4.3.5), we have

$$\delta_z = \begin{cases} -\frac{\pi\epsilon^2(1-2\epsilon)la_1^2 a_2 m^2}{(I_z+2m\epsilon l^2)^2} cos(\phi_2 - 2\phi_1) & \text{if } n = 2; \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $n = 2$ is the necessary condition for generating nonzero geometric phase under assumption (4.3.3). With this condition, $\phi_2 - 2\phi_1 = 2k\pi$, for $k = 0, \pm 1, \cdots$ gives the largest phase shift and $\phi_2 - 2\phi_1 = (2k + 1)\pi/2$, for $k = 0, \pm 1, \cdots$, gives zero phase shift.

We next formulate the optimal control problem for this particular mechanical system. For convenience, we re-scale θ by the factor $\frac{m\epsilon l}{(I_z+2m\epsilon l^2)^2}$. Then the third equation of (4.3.1) becomes

$$\dot{\theta} = x_1^2 \dot{x}_2 - x_2^2 \dot{x}_1.$$

The optimal control problem is to find control $u_1(\cdot)$ and $u_2(\cdot)$ to

$$\text{minimize} \int_0^1 (u_1^2 + u_2^2)dt \tag{4.3.6a}$$

subject to

$$\begin{cases} \dot{x}_1 = u_1 \\ \dot{x}_2 = u_2 \\ \dot{\theta} = x_1^2 u_2 - x_2^2 u_1 \end{cases} \tag{4.3.6b}$$

with given boundary conditions

$$x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \quad \theta(0) = \theta_0, \quad \theta(1) = \theta_1. \tag{4.3.6c}$$

Remark 4.3.1: It should be noted that the difference of $\theta(0)$ and $\theta(1)$ should not be too large because of the assumption on our simplified model. ∎

Since the optimal control problem (4.3.6) has fixed boundary conditions, one needs to check reachability for the control system (4.3.6b). Define two vector fields on $\Re^2 \times S^1$ by

$$g_1(q) = \begin{pmatrix} 1 \\ 0 \\ x_2^2 \end{pmatrix} \qquad g_2(q) = \begin{pmatrix} 0 \\ 1 \\ -x_1^2 \end{pmatrix}.$$

83

Then (4.3.6b) can be represented as

$$\dot{q} = g_1(q)u_1 + g_2(q)u_2.$$

It is easy to check that

$$[g_1, [g_1, g_2]](q) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

and $g_1(q), g_2(q), [g_1, [g_1, g_2]](q)$ are linearly independent for any $q = (x_1, x_2, \theta) \in \Re^2 \times S^1$. By Chow's theorem, we conclude that, for the system (4.3.6b), there is an open set about the point $q = (0, 0, \theta)$ (or any other point in $\Re^2 \times S^1$) such that any point in this set can be reached from q by a piecewise constant input (u_1, u_2).

Theorem 4.3.2: If $(x_1(\cdot), x_2(\cdot), \theta(\cdot))$ is an optimal trajectory with control $(u_1^*(\cdot), u_2^*(\cdot))$ for the optimal control problem given by (4.3.6), then there exists $\lambda(\cdot) = (\lambda_1(\cdot), \lambda_2(\cdot), \lambda_3(\cdot))^T$ on $[0, 1]$ satisfying the ordinary differential equations

$$\begin{cases} \dot{x}_1 = u_1^* \\ \dot{x}_2 = u_2^* \\ \dot{\theta} = x_1^2 u_2^* - x_2^2 u_1^* \end{cases} \qquad \begin{cases} \dot{\lambda}_1 = -2\lambda_3 x_1 u_2^* \\ \dot{\lambda}_2 = 2\lambda_3 x_2 u_1^* \\ \dot{\lambda}_3 = 0, \end{cases} \qquad (4.3.7a)$$

where

$$u_1^*(q) = -\frac{1}{2}(\lambda_1 - \lambda_3 x_2^2) \quad \text{and} \quad u_2^*(q) = \frac{1}{2}(\lambda_2 + \lambda_3 x_1^2) \qquad (4.3.7b)$$

with boundary conditions

$$x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \theta(0) = \theta_0, \theta(1) = \theta_1. \qquad (4.3.7c)$$

Moreover, the system (4.3.7) is completely integrable.

Proof:

The equations (4.3.7) can be derived easily from the Maximum Principle. The derivation is omitted here. We just prove solvability. From (4.3.7) one can get differential equations for the geodesics:

$$\begin{cases} \ddot{x}_1 - \lambda_3(x_1 + x_2)\dot{x}_2 = 0 \\ \ddot{x}_2 + \lambda_3(x_1 + x_2)\dot{x}_1 = 0 \\ \ddot{x}_3 + 2(x_2 - x_1)\dot{x}_1\dot{x}_2 + \lambda_3(x_1 + x_2)(x_1^2\dot{x}_1 + x_2^2\dot{x}_2) = 0 \end{cases} \qquad (4.3.8)$$

84

for some constant λ_3. To integrate (4.3.8), let $w = x_1 + x_2$ and $v = x_1 - x_2$. We have

$$\begin{cases} \ddot{v} = \lambda_3 w \dot{w} \\ \ddot{w} = -\lambda_3 w \dot{v}. \end{cases}$$

By integrating the first equation and substituting the result in second equation, we get, for some constant c,

$$\ddot{w} + \lambda_3 w (c + \frac{\lambda_3}{2} w^2) = 0,$$

which is the equation for a quartic oscillator, solvable by elliptic functions, i.e.

$$t = \int \frac{dw}{\sqrt{C_1 + aw^2 + bw^4}} + C_2$$

for $a = \frac{\lambda_3 c}{2}, b = \frac{(\lambda_3)^2}{8}$, where C_1 and C_2 are integral constants. Therefore, we conclude that if there exists an optimal solution $(q(t), u(t))$ to (4.3.6), it can be determined explicitly, i.e. the boundary value problem (4.3.6) is solvable. ∎

4.4 Three Dimensional System

Starting from this section, we assume that the system is free to move in three dimensional space. Again, we assume that the masses of the two oscillators are equal, i.e., $m_1 = m_2$. On the carrier a coordinate system $0\text{-}xyz$ is set such that the three axes are principal axes, i.e., the moment of inertia I_0 of the carrier can be represented as

$$I_0 = Diag(I_x, I_y, I_z).$$

Two oscillators are allowed to move on the carrier such that Q_1 and Q_2 satisfy

$$Q_1(t) = (x_1(t)cos(\psi_1), \ x_1(t)sin(\psi_1), \ l)^T$$

$$Q_2(t) = (x_2(t)cos(\psi_2), \ x_2(t)sin(\psi_2), \ -l)^T,$$

where $l > 0$ and ψ_i for $i = 1, 2$ are constants. This means that the two oscillators are restricted to move along their guide-ways which are parallel to the $0\text{-}xy$ plane and are at an equal distance (l) from the plane (see Figure 4.4.1). The configuration space now becomes $Q = \Re^2 \times SO(3)$ which will be coordinatized by (x_1, x_2, A), and the principal bundle is $\wp = (\Re^2 \times S\dot{O}(3), \Re^2, \pi, SO(3))$.
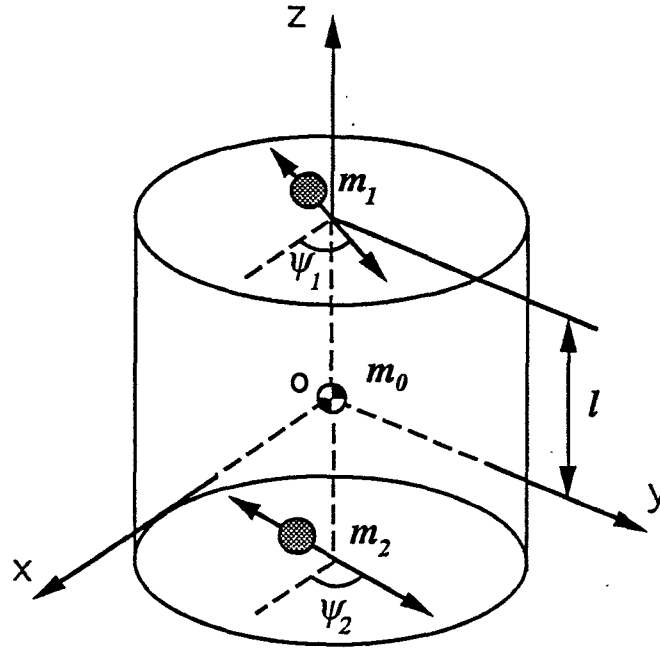
Figure 4.4.1 A rigid body with two oscillators: 3D system

In the above setting and under the condition $\mu = 0$, the angular velocity of the carrier in (4.2.14) now is of the form

$$\Omega = \Omega_1(x_1, x_2)\dot{x}_1 + \Omega_2(x_1, x_2)\dot{x}_2, \qquad (4.4.1)$$

where

$$\Omega_1 \triangleq \frac{1}{det(I_{lock})} \begin{pmatrix} \omega_{11} \\ \omega_{12} \\ \omega_{13} \end{pmatrix} \qquad \Omega_2 \triangleq \frac{1}{det(I_{lock})} \begin{pmatrix} \omega_{21} \\ \omega_{22} \\ \omega_{23} \end{pmatrix}$$

with $det(I_{lock})$ and ω_{ij}, for $i = 1, 2$ and $j = 1, 2, 3$, being polynomials of x_1 and x_2. And, the local connection form for the mechanical connection on principal bundle $\wp$ is

$$\omega = -\Omega_1(x_1, x_2)dx_1 - \Omega_2(x_1, x_2)dx_2. \qquad (4.4.2)$$

Although the above restriction on the motion of the two oscillators simplifies the bundle structure of the system, the expression of the connection form is still very complicated. In the rest of this section we will only consider some problems with special ψ_1 and ψ_2.

An interesting question is how to choose the constant parameters so that the above three dimensional system reduces to the planar system discussed in preceding section.

A natural guess may be that

$$\psi_1 = \psi_2 \stackrel{\triangle}{=} \psi. \tag{4.4.3}$$

But this is not the only condition. In fact, when (4.4.3) satisfies, ω_{13} and ω_{23} have simple expressions:

$$\omega_{13} = -\omega_{23} = \frac{1}{2}\epsilon^2(I_y - I_x)m^2 l^2 sin(2\psi)(x_2 - x_1).$$

Thus, if $I_x = I_y$ or $\psi = 0$ or $\psi = \frac{\pi}{2}$, then $\omega_{13} = \omega_{23} = 0$, i.e., the rigid body will only move (rotate) about the axis perpendicular to the plane formed by the guide-ways of two oscillators. Otherwise, in general, the parallel motion of the two oscillators will cause the rigid body to drift about z axis. In other words, when $\psi_1 = \psi_2$ holds, a sufficient condition for planar drift is that the carrier has axial symmetry about z-axis, or that two point masses move along the lines which are parallel with the same principal axis (0-x or 0-y) of the carrier.

Explicitly, if $\psi = 0$, i.e., both oscillators move parallel to principal axis 0-x, the local connection form is

$$\omega = (0, \ -\frac{\epsilon m l(dx_1 - dx_2)}{m\epsilon(\epsilon - 1)x_2^2 + 2\epsilon^2 m x_1 x_2 + \epsilon m(\epsilon - 1)x_1^2 - 2\epsilon m l^2 - I_y}, \ 0)^T.$$

If $\psi = \frac{\pi}{2}$, i.e., both oscillators move parallel to principal axis 0-y, the local connection form is

$$\omega = (\frac{\epsilon m l(dx_1 - dx_2)}{m\epsilon(\epsilon - 1)x_2^2 + 2\epsilon^2 m x_1 x_2 + \epsilon m(\epsilon - 1)x_1^2 - 2\epsilon m l^2 - I_x}, \ 0, \ 0)^T.$$

The nonzero terms in above ω's are the same as ω in (4.3.2) (up to a choice of the coordinates system).

From the above discussion, it is apparent that if one is interested in full three dimensional motion of the carrier body, some skewness in the directions of particle motions would be necessary. In the following, we set $\psi_1 = 0$ and $\psi_2 = \frac{\pi}{2}$. With this setting, the constraint on the motion of two point masses becomes

$$Q_1 = (x_1(t), 0, -l)^T \quad \text{and} \quad Q_2 = (0, x_2(t), l)^T.$$

The body angular velocity Ω can still be given by

$$\Omega = \Omega_1(x_1, x_2)\dot{x}_1 + \Omega_2(x_1, x_2)\dot{x}_2$$

for Ω_1 and Ω_2 given in (4.4.1). The explicit expressions of the components of Ω_i, ω_{ij}, are still complicated. We show them in Appendix 4.A.

We now show that the kinematic control system corresponding to (4.1.17) with $\mu = 0$ is controllable in a neighborhood of $(0, 0, A)$ for any $A \in SO(3)$. From (4.2.12) with $\mu = 0$, the kinematic control system is of the form

$$\begin{cases} \dot{x}_1 = u_1 \\ \dot{x}_2 = u_2 \\ \dot{A} = A(\widehat{\Omega}_1 u_1 + \widehat{\Omega}_2 u_2). \end{cases} \tag{4.4.4}$$

Let $x = (x_1, x_2)$ and $q = (x, A)$ be a point in Q. Equation (4.4.4) can be represented as

$$\dot{q} = X_1(q)u_1 + X_2(q)u_2, \tag{4.4.4'}$$

where $X_1(q) = ((1, 0), A\widehat{\Omega}_1(x))$ and $X_2(q) = ((0, 1), A\widehat{\Omega}_2(x))$.

Again, we can use the rank condition in Chow's theorem [9] to check the controllability of the system. To this end, we need a formula to compute the Lie bracket of vector fields X_1 and X_2 on $\Re^2 \times SO(3)$, where X_i is represented as, in general,

$$X_i(x, A) = (F_i(x), A\widehat{G}_i(x)) \qquad i = 1, 2 \tag{4.4.5}$$

for smooth mappings $F_i : \Re^2 \to \Re^2$ and $G_i : \Re^2 \to \Re^3$.

Remark 4.4.1: The classical methods to express Equations (4.4.4), as well as vector fields in (4.4.5) is to introduce Euler angles or other local coordinates, which are very important for numerical calculation. However, such expressions are usually very complicated and very hard to use for qualitative analysis. In the following analysis, we will use an intrinsic method, namely, keeping the differential equation (4.4.4) in matrix form and using the orthogonal property for the matrices only. As we will see later, this method will give us a clean expression for the necessary conditions for the optimal control problem and make reduction of the necessary condition possible. For other references on the same treatment of problems on modeling and control of mechanical systems on $SO(3)$, see [35,49]. ∎

Let X be a smooth vector field on a smooth manifold P. Let $\phi(\cdot, p) : \Re \to P : t \mapsto \phi(t, p)$ to be *the tangent curve of X at p,* i.e., it is a smooth curve in P such

that $\phi(0, p) = p$ and $\frac{d}{dt}|_{t=0}\phi(t, p) = X(p)$. All of such tangent curves at p forms an equivalent class, which is denoted by $[\![\phi(t, p)]\!]$. Then, the tangent vector at p can be represented as

$$X(p) = [\![\phi(t, p)]\!]$$

and the Lie derivative of a smooth function f on P along vector field X, $L_X f(p) = df(p) \cdot X(p)$, can be computed by using the equivalence class of curves, i.e.,

$$L_X f(p) = [\![f \circ \phi(t, p)]\!]. \tag{4.4.6}$$

This expression is well defined since $f \circ \phi(\cdot, p) : \Re \to \Re$ for given point p on P. We then have the following result.

Lemma 4.4.2: Let X_1 and X_2 be two smooth vector fields on P, and ϕ_1 and ϕ_2 be any corresponding tangent curves at p. Then

$$[\, X_1, X_2 \,](f)(p) = \frac{d}{dt}\bigg|_{t=0} (X_2(\phi_1(t, p)) - X_1(\phi_2(t, p)))(f)(p) \tag{4.4.7}$$

for any $p \in P$ and any smooth function f on P.

Proof: With notation (4.4.6), in local coordinates, we have

$$
\begin{aligned}
L_{X_1} L_{X_2} f(p) &= [\![L_{X_2} f \circ \phi_1(t, p)]\!] \\
&= [\![[\![f \circ \phi_2(s, \phi_1(t, p))]\!]]\!] \\
&= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} f(\phi_2(s, \phi_1(t, p))) \\
&= \frac{d}{dt}\bigg|_{t=0} Df(\phi_1(t, p)) \cdot X_2(\phi_1(t, p)) \\
&= (DDf(p) \cdot X_1(p)) \cdot X_2(p) + Df(p) \cdot \frac{d}{dt}\bigg|_{t=0} X_2(\phi_1(t, p)).
\end{aligned}
$$

Since $(DDf(p) \cdot X_1(p)) \cdot X_2(p) = (DDf(p) \cdot X_2(p)) \cdot X_1(p)$, the Lie bracket of vector fields X_1 and X_2 is

$$
\begin{aligned}
[\, X_1, X_2 \,](f)(p) &= L_{X_2} L_{X_1}(f)(p) - L_{X_1} L_{X_2}(f)(p) \\
&= \frac{d}{dt}\bigg|_{t=0} (X_2(\phi_1(t)) - X_1(\phi_2(t)))(f)(p)
\end{aligned}
$$

for any $p \in P$ and any smooth function f on P. ∎

Now we use (4.4.7) to compute the Lie bracket of two vector fields on $\Re^2 \times SO(3)$ given in (4.4.5).

Proposition 4.4.3: Given two vector fields X_1 and X_2 on $\Re^2 \times SO(3)$ as shown in (4.4.5), the Lie bracket of X_1 and X_2 is given by

$$[\,X_1, X_2\,](x, A) = (\frac{\partial F_2}{\partial x}F_1 - \frac{\partial F_1}{\partial x}F_2, \quad A[G_1 \times G_2 + \frac{\partial G_2}{\partial x}F_1 - \frac{\partial G_1}{\partial x}\widehat{F_2}]\,) \qquad (4.4.8)$$

for any point $(x, A) \in \Re^2 \times SO(3)$.

Proof: It is clear that the tangent curve of vector fileds X_i in (4.4.5) at (x, A) can be represented as

$$\phi_i(\tau) = (x + \tau F_i(x), Aexp(\tau \widehat{G}_i(x))).$$

Then, by (4.4.7),

$$
\begin{aligned}
[\,X_1, X_2\,](x, A) &= \frac{d}{d\tau}\Big|_{\tau=0} (F_2(x + \tau F_1) - F_1(x + \tau F_2), \\
&\qquad Aexp(\tau \widehat{G}_1)\widehat{G}_2(x + \tau F_1) - Aexp(\tau \widehat{G}_2)\widehat{G}_1(x + \tau F_2)) \\
&= (\frac{\partial F_2}{\partial x}F_1 - \frac{\partial F_1}{\partial x}F_2, A(\widehat{G}_1\widehat{G}_2 - \widehat{G}_2\widehat{G}_1 + [\frac{\partial G_2}{\partial x}F_1 - \frac{\partial G_1}{\partial x}\widehat{F_2}]\,)) \\
&= (\frac{\partial F_2}{\partial x}F_1 - \frac{\partial F_1}{\partial x}F_2, \quad A[G_1 \times G_2 + \frac{\partial G_2}{\partial x}F_1 - \frac{\partial G_1}{\partial x}\widehat{F_2}]\,). \qquad \blacksquare
\end{aligned}
$$

Theorem 4.4.4: For system (4.4.4), there is an open set about $(0, 0, A)$ for any $A \in SO(3)$ such that any point in this set can be reached from $(0, 0, A)$ by a piecewise constant input (u_1, u_2).

Proof: Let

$$X_3(x, A) \stackrel{\triangle}{=} [\,X_1, X_2\,](x, A) = (F_3(x), A\widehat{G}_3(x)),$$

$$X_4(x, A) \stackrel{\triangle}{=} [\,X_1, X_3\,](x, A) = (F_4(x), A\widehat{G}_4(x)),$$

$$X_5(x, A) \stackrel{\triangle}{=} [\,X_2, X_3\,](x, A) = (F_5(x), A\widehat{G}_5(x)),$$

where F_i and G_i, $i = 1, 2, 3$ are computed by using (4.4.7). It is easy to see that

$$F_i = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{for} \quad i = 3, 4, 5.$$

Using symbolic calculation (say, MACSYMA), one can check that in general,

$$det(G_3, G_4, G_5)|_{x=(0,0)} \neq 0.$$

90

Since $T_{(x,A)}(\Re^2 \times SO(3)) \simeq \Re^5$ and the vector fields shown in (4.4.4)', which generate the smallest involutive distribution, have special form, namely, $F_1(x) = (1,0)^T$ and $F_2(x) = (0,1)^T$, the above equations are sufficient to show that vector fields X_i for $i = 1, ..., 5$ are independent. Consequently, the control system given in (4.4.4) is controllable by Chow's theorem. ∎

We now turn to the optimal control problem. Corresponding to (P2) in Subsection 4.1.2, the goal here is to find $u_1(\cdot)$ and $u_2(\cdot)$ to

$$\text{minimize} \int_0^1 (u_1^2 + u_2^2)dt \tag{4.4.9a}$$

subject to

$$\begin{cases} \dot{x}_1 = u_1 \\ \dot{x}_2 = u_2 \\ \dot{A} = A(\widehat{\Omega}_1 u_1 + \widehat{\Omega}_2 u_2) \end{cases} \tag{4.4.9b}$$

for given boundary conditions

$$x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \ A(0) = A_0 \in SO(3), \ A(1) = A_1 \in SO(3). \tag{4.4.9c}$$

The necessary conditions for the above problem is given in the following theorem.

Theorem 4.4.5: If $(x(\cdot), A(\cdot))$ is an optimal trajectory with controls $(u_1^*(\cdot), u_2^*(\cdot))$ for the optimal control problem given by (4.4.9), then there exist $\mu(\cdot) = (\mu_1(\cdot), \mu_2(\cdot))$ and $\lambda(\cdot)$ on $[0, 1]$ satisfying the ordinary differential equations

$$\begin{cases} \dot{x}_1 = u_1^* \\ \dot{x}_2 = u_2^* \\ \dot{A} = A(\widehat{\Omega}_1 u_1^* + \widehat{\Omega}_2 u_2^*) \end{cases} \quad \begin{cases} \dot{\mu}_1 = -\lambda^T(\dfrac{\partial \Omega_1}{\partial x_1} u_1^* + \dfrac{\partial \Omega_2}{\partial x_1} u_2^*) \\ \dot{\mu}_2 = -\lambda^T(\dfrac{\partial \Omega_1}{\partial x_2} u_1^* + \dfrac{\partial \Omega_2}{\partial x_2} u_2^*) \\ \dot{\lambda} = \lambda \times (\Omega_1 u_1^* + \Omega_2 u_2^*), \end{cases} \tag{4.4.10a}$$

where

$$\begin{cases} u_1^*(x, A, \mu, \lambda) = -\dfrac{1}{2}(\mu_1 + \lambda^T \Omega_1) \\ u_2^*(x, A, \mu, \lambda) = -\dfrac{1}{2}(\mu_2 + \lambda^T \Omega_2) \end{cases} \tag{4.4.10b}$$

with boundary conditions

$$x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, A(0) = A_0 \in SO(3), A(1) = A_1 \in SO(3). \tag{4.4.10c}$$

Remark 4.4.6: Note that the above theorem does not depend on the expressions of vectors Ω_1 and Ω_2 and the dimension of the shape space (now it is two) is not an important factor as we will see in the following proofs. Therefore, the above theorem is applicable to a class of mechanical systems with product principal bundle $(\Re^m \times SO(3), \Re^m, \pi, SO(3))$. ■

In the following we use two methods to prove Theorem 4.4.5: perturbation method (in variational principle) and geometric method.

Proof of Theorem 4.4.5 (Perturbation Method):

Applying Lagrange multiplier, the optimal control problem is equivalent to finding control $(u_1(\cdot), u_2(\cdot))$ to minimize the cost function

$$J(u_1(\cdot), u_2(\cdot)) = \int_0^1 (u_1^2 + u_2^2 + \mu_1(u_1 - \dot{x}_1) + \mu_2(u_2 - \dot{x}_2) + \langle \Lambda_A, A(\widehat{\Omega}_1 u_1 + \widehat{\Omega}_2 u_2) - \dot{A} \rangle) dt,$$

where $(\mu_1, \mu_2) \in T^*_{(x_1, x_2)} \Re^2 (\simeq \Re^2)$, and $\Lambda_A \in T^*_A SO(3)$, which are Lagrange multipliers, $\langle \cdot, \cdot \rangle$ is the natural pairing between $T^* SO(3)$ and $T SO(3)$ defined in (4.2.8). Since $\Lambda_A \in T^*_A SO(3)$, it is of the form $\Lambda_A = A \widehat{\lambda}$ for some $\lambda \in \Re^3$. Then the cost function becomes

$$J(u_1(\cdot), u_2(\cdot)) = \int_0^1 \widetilde{H}(x_1, \dot{x}_1, x_2, \dot{x}_2, A, \dot{A}, u_1, u_2; \mu_1, \mu_2, \lambda) dt,$$

where

$$\widetilde{H} = u_1^2 + u_2{}^2 + \mu_1(u_1 - \dot{x}_1) + \mu_2(u_2 - \dot{x}_2) + \langle \widehat{\lambda}, \widehat{\Omega}_1 u_1 + \widehat{\Omega}_2 u_2 \rangle - \langle A\widehat{\lambda}, \dot{A} \rangle \quad (4.4.11)$$

which is a mapping from $\Re^2 \times T^*(\Re^2 \times SO(3)) \times T(\Re^2 \times SO(3))$ to $\Re$. Due to the presence of the last term of (4.4.11) we cannot directly apply formulae for the adjoint equations and optimal control as appearing in most text books. However, we still can use the idea of perturbation in variational principle for our problem.

First, consider the perturbation of function (4.4.11). Let

$$\mathcal{C}_c = \{(u_1(t), u_2(t)), t \in [0, 1]\}$$

be the optimal curve in control space and

$$\mathcal{C}_s = \{(x_1(t), x_2(t), A(t)), t \in [0, 1]\}$$

the corresponding optimal curve in configuration space $\Re^2 \times SO(3)$. Let

$$\mathcal{C}_{cp} = \{(u_1(t) + \epsilon\phi_{u_1}(t), u_2(t) + \epsilon\phi_{u_2}(t)), t \in [0,1]\}$$

be the perturbed curve of $\mathcal{C}_c$, where $\epsilon \in \Re$, $\phi_{u_1}(\cdot)$ and $\phi_{u_2}(\cdot)$ are smooth curve in $\Re^1$. The corresponding perturbed curve of $\mathcal{C}_s$ is presented by

$$\mathcal{C}_{sp} = \{(x_1(t) + \epsilon\phi_{x_1}(t), x_2(t) + \epsilon\phi_{x_2}(t), Ae^{\epsilon\widehat{\phi}_A(t)}), t \in [0,1]\},$$

where $\phi_{x_1}(\cdot)$ and $\phi_{x_2}(\cdot)$ are smooth curves in $\Re^1$ satisfying $\phi_{x_1}(0) = \phi_{x_1}(1) = \phi_{x_2}(0) = \phi_{x_2}(1) = 0$ and $\phi_A(\cdot)$ is a smooth curve in $\Re^3$ satisfying $\phi_A(0) = \phi_A(1) = \mathbf{0}$. It should be noted that, since $exp : T_eG \to G$ for any Lie group G is locally diffeomorphism, $Ae^{\epsilon\widehat{\phi}_A} \in SO(3)$ is qualified as a perturbation of A. Now the integrand of perturbed cost function J_ϵ is

$$
\begin{aligned}
\tilde{H}_\epsilon =& (u_1 + \epsilon\phi_{u_1})^2 + (u_2 + \epsilon\phi_{u_2})^2 \\
& + \mu_1(u_1 + \epsilon\phi_{u_1} - \dot{x}_1 - \epsilon\dot{\phi}_{x_1}) + \mu_2(u_2 + \epsilon\phi_{u_2} - \dot{x}_2 - \epsilon\dot{\phi}_{x_2}) \\
& + \lambda^T[\Omega_1(x_1 + \epsilon\phi_{x_1}, x_2 + \epsilon\phi_{x_2})(u_1 + \epsilon\phi_{u_1}) \\
& + \Omega_2(x_1 + \epsilon\phi_{x_1}, x_2 + \epsilon\phi_{x_2})(u_2 + \epsilon\phi_{u_2})] \\
& - \langle Ae^{\epsilon\widehat{\phi}_A}\widehat{\lambda}, \dot{A}e^{\epsilon\widehat{\phi}_A} + A\frac{d}{dt}(e^{\epsilon\widehat{\phi}_A})\rangle.
\end{aligned}
$$

By letting the derivative of J_ϵ with respect to ϵ at $\epsilon = 0$ be zero, i.e., $\frac{dJ_\epsilon}{d\epsilon}|_{\epsilon=0} = 0$, we have

$$
\begin{aligned}
0 = \int_0^1 &[2u_1\phi_{u_1} + 2u_2\phi_{u_2} + \mu_1(\phi_{u_1} - \dot{\phi}_{x_1}) + \mu_2(\phi_{u_2} - \dot{\phi}_{x_2}) \\
& + \lambda^T((\frac{\partial\Omega_1}{\partial x_1}\phi_{x_1} + \frac{\partial\Omega_1}{\partial x_2}\phi_{x_2})u_1 + \Omega_1\phi_{u_1} + (\frac{\partial\Omega_2}{\partial x_1}\phi_{x_1} + \frac{\partial\Omega_2}{\partial x_2}\phi_{x_2})u_2 + \Omega_2\phi_{u_2}) \\
& - \langle A\widehat{\phi}_A\widehat{\lambda}, \dot{A}\rangle - \langle A\widehat{\lambda}, \dot{A}\widehat{\phi}_A + A\dot{\widehat{\phi}}_A\rangle]dt.
\end{aligned}
\tag{4.4.12}
$$

The last two terms in the integrand of (4.4.12) can be re-arranged as

$$
\begin{aligned}
\langle A\widehat{\phi}_A\widehat{\lambda}, \dot{A}\rangle &+ \langle A\widehat{\lambda}, \dot{A}\widehat{\phi}_A\rangle + \langle A\widehat{\lambda}, A\dot{\widehat{\phi}}_A\rangle \\
&= \frac{1}{2}tr(A\widehat{\phi}_A\widehat{\lambda}\dot{A}^T + A\widehat{\lambda}\widehat{\phi}_A^T\dot{A}^T + \widehat{\lambda}^T\dot{\widehat{\phi}}_A) \\
&= \frac{1}{2}tr(-\widehat{\phi}_A\widehat{\lambda}\widehat{\Omega} + \widehat{\lambda}\widehat{\phi}_A\widehat{\Omega} - \dot{\widehat{\lambda}}^T\widehat{\phi}_A) + \frac{d}{dt}(\lambda \cdot \phi_A) \\
&= -(\dot{\lambda} - \lambda \times \Omega) \cdot \phi_A + \frac{d}{dt}(\lambda \cdot \phi_A),
\end{aligned}
$$

where $\Omega = (A^T \dot{A})\check{} = \Omega_1 u_1 + \Omega_2 u_2$. Here we used a fact,

$$\widehat{\Omega}\widehat{\lambda} - \widehat{\lambda}\widehat{\Omega} = [\,\widehat{\Omega}, \widehat{\lambda}\,] = \widehat{\Omega \times \lambda}. \tag{4.4.13}$$

Also re-arranging the other terms in (4.4.12), we get

$$\begin{aligned}
0 = \int_0^1 [&-\frac{d}{dt}(\lambda \cdot \phi_A + \mu_1 \phi_{x_1} + \mu_2 \phi_{x_2}) \\
&+ (2u_1 + \mu_1 + \lambda^T \Omega_1)\phi_{u_1} \\
&+ (2u_2 + \mu_2 + \lambda^T \Omega_2)\phi_{u_2} \\
&+ (\dot{\mu}_1 + \lambda^T(\frac{\partial \Omega_1}{\partial x_1}u_1 + \frac{\partial \Omega_2}{\partial x_1}u_2))\phi_{x_1} \\
&+ (\dot{\mu}_2 + \lambda^T(\frac{\partial \Omega_1}{\partial x_2}u_1 + \frac{\partial \Omega_2}{\partial x_2}u_2))\phi_{x_2} \\
&+ (\dot{\lambda} - \lambda \times \Omega) \cdot \phi_A]dt.
\end{aligned}$$

Now, by the conditions on ϕ_{x_1}, ϕ_{x_2} and ϕ_A at $t = 0$ and 1, the first term of above integration is zero. In addition, we can choose functions $\mu_1(\cdot), \mu_2(\cdot)$ and $\lambda(\cdot)$ on $[0, 1]$ such that the coefficients of ϕ_{x_1}, ϕ_{x_2} and ϕ_A to be zeros; this gives the last three equations of (4.4.10a). Finally, since the system is controllable at $(0, 0, A)$ (cf. Theorem 4.4.4), the coefficients of ϕ_{u_1} and ϕ_{u_2} are also zero; this gives u_1^* and u_2^* in (4.4.10b).

$$\triangle$$

Proof of Theorem 4.4.5 (Geometric Method):

Here, we consider a slightly more general form of the problem (4.4.9). The optimal control problem now is to determine control $(u_1(\cdot), u_2(\cdot))$ to

$$\text{minimize} \int_0^1 (u_1^2 + u_2^2)dt \tag{4.4.14a}$$

subject to

$$\begin{cases} \dot{x} = f_1(x)u_1 + f_2(x)u_2 \\ \dot{A} = A(\widehat{\Omega}_1(x)u_1 + \widehat{\Omega}_2(x)u_2) \end{cases} \tag{4.4.14b}$$

with boundary conditions

$$(x(0), A(0)) = (x_0, A_0) \quad \text{and} \quad (x(1), A(1)) = (x_1, A_1), \tag{4.4.14c}$$

where $(x(t), A(t)) \in \Re^2 \times SO(3), \forall t \in [0, 1]$, $f_i : \Re^2 \to \Re^2$, $\Omega_i : \Re^2 \to \Re^3$ for $i = 1, 2$ and $u(\cdot) = (u_1(\cdot), u_2(\cdot)) : [0, 1] \to \Re^2$ is a piecewise smooth function. Applying the Maximum Principle to this problem, we have the following result.

94

Denoting by $z = ((x, A), (\mu, \Lambda_A)) = ((x, A), (\mu, A\widehat{\lambda})) \simeq ((x, A), (\mu, \lambda))$ a point in $T^*(\Re^2 \times SO(3))$, we define *pseudo-Hamiltonian*, $\mathcal{H} : T^*(\Re^2 \times SO(3)) \times \Re^2 \to \Re$, by

$$\mathcal{H}(z, u) \triangleq\; <u, u> + <\mu, f_1(x)u_1 + f_2(x)u_2> + \langle \Lambda_A, A(\widehat{\Omega}_1(x)u_1 + \widehat{\Omega}_2(x)u_2) \rangle$$

$$=<u, u> + <\mu, f_1(x)u_1 + f_2(x)u_2> + <\lambda, \Omega_1(x)u_1 + \Omega_2 u_2>.$$

Define Hamiltonian

$$H(z) \triangleq \min_{u \in \Re^2} \mathcal{H}(z, u).$$

Since the control space is unbounded, one can find the functions $u^*(z) = (u_1^*(z), u_2^*(z))$ on $T^*(\Re^2 \times SO(3))$ with

$$\begin{cases} u_1^*(z) = -\dfrac{1}{2}(\mu^T f_1(x) + \lambda^T \Omega_1(x)) \\[2mm] u_2^*(z) = -\dfrac{1}{2}(\mu^T f_2(x) + \lambda^T \Omega_2(x)) \end{cases} \tag{4.4.15}$$

such that

$$H(z) = \mathcal{H}(z, u^*(z)) = -\frac{1}{4} \sum_{i=1}^{2} (<\mu, f_i> + <\lambda, \Omega_i>)^2. \tag{4.4.16}$$

From the Maximum Principle, we know that the existence of optimal control for problem (4.4.14) implies that there is a solution to the following system on $T^*(\Re^2 \times SO(3))$:

$$\dot{z} = X_H(z) \tag{4.4.17}$$

with

$$\tau(z(0)) = (x_0, A_0) \quad \text{and} \quad \tau(z(1)) = (x_1, A_1),$$

where X_H is the Hamiltonian vector filed with respect to the Hamiltonian (4.4.16) and $\tau : T^*(\Re^2 \times SO(3)) \to \Re^2 \times SO(3)$ is the canonical projection. Our goal now is to determine the vector field X_H on $T^*(\Re^2 \times SO(3))$.

Recall that, for given n-dimensional smooth manifold Q, its cotangent bundle T^*Q has a canonical symplectic form Ω_0. Given a Hamiltonian H on T^*Q, the corresponding Hamiltonian vector field X_H on T^*Q is defined by

$$\Omega_0(X_H, Y) = H(Y) \tag{4.4.18}$$

for any $Y \in \mathfrak{X}(T^*Q)$. The local expression of (4.4.18) can be given as (cf. Theorem 3.2.10 in [1])

$$\Omega_0(x, \alpha)((x, \alpha, e_1, \beta_1), (x, \alpha, e_2, \beta_2)) = \langle \beta_2, e_1 \rangle - \langle \beta_1, e_2 \rangle \tag{4.4.19}$$

for $(x, \alpha) \in T^*Q$ and $(e_i, \beta_i) \in T_{(x,\alpha)}T^*Q$, $i = 1, 2$. In our problem, $Q = \Re^2 \times SO(3)$.
For the Hamiltonian given in (4.4.16), the corresponding Hamiltonian vector field X_H
in (4.4.17) will be determined by (4.4.18).

Let

$$y(t) = ((x + tv, Ae^{t\widehat{\alpha}}), (\mu + t\phi, Ae^{t\widehat{\alpha}}(\widehat{\lambda} + t\widehat{\beta})))$$

be an integral curve on $T^*(\Re^2 \times SO(3))$ at $z = ((x, A), (\mu, A\widehat{\lambda}))$ for any $v, \phi \in \Re^2$,
$\alpha, \beta \in \Re^3$. Then its tangent vector at z is given by

$$Y(z) = ((x, A), (\mu, A\widehat{\lambda}), (v, A\widehat{\alpha}), (\phi, A(\widehat{\alpha}\widehat{\lambda} + \widehat{\beta}))).$$

Now the right-hand-side of (4.4.18) can be calculated as

$$H(Y) = \frac{d}{dt}\Big|_{t=0} H(y(t))$$

$$= -\frac{1}{4}\frac{d}{dt}\Big|_{t=0} \sum_{i=1}^{2}(< \mu + t\phi, f_i(x + tv) > + < \lambda + t\beta, \Omega_i(x + tv) >)^2$$

$$= -\frac{1}{2}\sum_{i=1}^{2}(< \mu, f_i > + < \lambda, \Omega_i >) \cdot$$

$$(< \phi, f_i > + < \mu, Df_i \cdot v > + < \beta, \Omega_i > + < \lambda, D\Omega_i \cdot v >)$$

$$= < \phi, f_1 u_1^* + f_2 u_2^* > + < \beta, \Omega_1 u_1^* + \Omega_2 u_2^* >$$

$$+ < \mu, (Df_1 u_1^* + Df_2 u_2^*) \cdot v > + < \lambda, (D\Omega_1 u_1^* + D\Omega_2 u_2^*) \cdot v >, (4.4.20)$$

where u_i^* is given in (4.4.15) and

$$Df_i = \left(\begin{matrix} \frac{\partial f_i}{\partial x_1} & \frac{\partial f_i}{\partial x_2} \end{matrix}\right) \quad \text{and} \quad D\Omega_i = \left(\begin{matrix} \frac{\partial \Omega_i}{\partial x_1} & \frac{\partial \Omega_i}{\partial x_2} \end{matrix}\right)$$

for $i = 1, 2$ are 2×2 and 3×2 matrices, respectively.

On the other hand, let

$$X_H(z) = ((x, A), (\mu, A\widehat{\lambda}), (w, A\widehat{\epsilon}), (\eta, A(\widehat{\epsilon}\widehat{\lambda} + \widehat{\delta})))$$

for some vectors $w, \eta \in \Re^2$ and $\epsilon, \delta \in \Re^3$ which will be determined. Applying (4.4.19),
we have

$$\Omega_0(z)(X_H, Y) = \langle (\phi, A(\widehat{\alpha}\widehat{\lambda} + \widehat{\beta})), (w, A\widehat{\epsilon}) \rangle - \langle (\eta, A(\widehat{\epsilon}\widehat{\lambda} + \widehat{\delta})), (v, A\widehat{\alpha}) \rangle$$

$$= < \phi, w > + \langle A(\widehat{\alpha}\widehat{\lambda} + \widehat{\beta}), A\widehat{\epsilon} \rangle - < \eta, v > - \langle A(\widehat{\epsilon}\widehat{\lambda} + \widehat{\delta}), A\widehat{\alpha} \rangle$$

$$= < \phi, w > - < \eta, v > + \frac{1}{2}tr(\widehat{\beta}^T\widehat{\epsilon} - \widehat{\delta}^F\widehat{\alpha} + \widehat{\lambda}^T[\widehat{\epsilon}, \widehat{\alpha}])$$

$$= < \phi, w > - < \eta, v > + < \beta, \epsilon > - \frac{1}{2}tr([\widehat{\epsilon}, \widehat{\lambda}] + \widehat{\delta})\widehat{\alpha}). \quad (4.4.21)$$

96

In order to have equation (4.4.18), we need to make (4.4.20) equal to (4.4.21). This leads to following choice of w, ϵ, η and δ.

$$w = f_1 u_1^* + f_2 u_2^*$$

$$\eta = -(Df_1 u_1^* + Df_2 u_2^*)^T \mu - (D\Omega_1 u_1^* + D\Omega_2 u_2^*)^T \lambda$$

$$\epsilon = \Omega_1 u_1^* + \Omega_2 u_2^* \tag{4.4.22}$$

$$\delta = \lambda \times (\Omega_1 u_1^* + \Omega_2 u_2^*).$$

With above equations, the vector field X_H can be completely determined.

In summary, the differential equation (4.4.17) now has the following form.

$$\begin{cases} \dot{x} = f_1(x)u_1^* + f_2(x)u_2^* \\ \dot{A} = A(\widehat{\Omega}_1(x)u_1^* + \widehat{\Omega}_2(x)u_2^*) \\ \dot{\mu} = -(Df_1(x)u_1^* + Df_2(x)u_2^*)^T \mu - (D\Omega_1(x)u_1^* + D\Omega_2(x)u_2^*)^T \lambda \\ \dot{\lambda} = \lambda \times (\Omega_1(x)u_1^* + \Omega_2(x)u_2^*) \end{cases} \tag{4.4.23a}$$

where

$$\begin{cases} u_1^* = -\dfrac{1}{2}(\mu^T f_1(x) + \lambda^T \Omega_1(x)) \\ u_2^* = -\dfrac{1}{2}(\mu^T f_2(x) + \lambda^T \Omega_2(x)). \end{cases} \tag{4.4.23b}$$

When $f_1 = (1,0)^T$ and $f_2 = (0,1)^T$, (4.4.23) leads to (4.4.10). ∎

Remark 4.4.7: From the second proof of Theorem 4.4.5, we can see that the geometric treatment of the optimal control problem allows us to define a Hamiltonian vector field on the manifold $T^*(\Re^2 \times SO(3))$. The solution of the optimal control problem will correspond to a trajectory of this vector field. Of course, in general, it is almost impossible to find explicit solutions although we did find it for planar system as shown in Section 4.3. However, as we will show in the next section, identification of symmetries in such a Hamiltonian system will allow a reduction in the order of the system by applying symplectic or Poisson reduction theory which we reviewed in Chapter II. ∎

4.5 Symmetry and Reduction

Recall that the manifold $P = T^*(\Re^2 \times SO(3))$, parameterized by $z = (x, A, \mu, A\widehat{\lambda})$, is symplectic. Hence, it is also Poisson (cf. Section 2.1). One can verify that the Poisson

bracket of smooth functions F_1 and F_2 on P is given by

$$\{F_1, F_2\}_P(z) = \frac{\partial F_1}{\partial x} \cdot \frac{\partial F_2}{\partial \mu} - \frac{\partial F_2}{\partial x} \cdot \frac{\partial F_1}{\partial \mu}$$
$$+ \langle D_A F_1, D_{\widehat{A\lambda}} F_2 \rangle - \langle D_A F_2, D_{\widehat{A\lambda}} F_1 \rangle. \tag{4.5.1}$$

In the preceding section, we have shown that, for Hamiltonian

$$H(z) = -\frac{1}{4} \sum_{i=1}^{2} (\mu_i + < \lambda, \Omega_i >)^2, \tag{4.5.2}$$

the Hamiltonian vector field is

$$X_H(z) = \begin{pmatrix} u_1^* \\ u_2^* \\ A(\widehat{\Omega}_1 u_1^* + \widehat{\Omega}_2 u_2^*) \\ -\lambda^T(\frac{\partial \Omega_1}{\partial x_1} u_1^* + \frac{\partial \Omega_2}{\partial x_1} u_2^*) \\ -\lambda^T(\frac{\partial \Omega_1}{\partial x_2} u_1^* + \frac{\partial \Omega_2}{\partial x_2} u_2^*) \\ \lambda \times (\Omega_1 u_1^* + \Omega_2 u_2^*) \end{pmatrix} \tag{4.5.3}$$

where

$$u_1^*(z) = -\frac{1}{2}(\mu_1 + \lambda^T \Omega_1)$$

$$u_2^*(z) = -\frac{1}{2}(\mu_2 + \lambda^T \Omega_2).$$

Consider the action of $SO(3)$ on $\Re^2 \times SO(3)$

$$\Phi : SO(3) \times (\Re^2 \times SO(3)) \to \Re^2 \times SO(3)$$

$$(B, (x, A)) \mapsto (x, BA)$$

and its cotangent lift on $T^*(\Re^2 \times SO(3))$

$$\Phi^{T^*} : SO(3) \times T^*(\Re^2 \times SO(3)) \to T^*(\Re^2 \times SO(3))$$

$$(B, (x, A, \mu, A\widehat{\lambda})) \mapsto (x, BA, \mu, BA\widehat{\lambda}).$$

The quotient space $\widetilde{P} = T^*(\Re^2 \times SO(3))/SO(3) \simeq \Re^2 \times \Re^2 \times \Re^3$ can be coordinatized by $\widetilde{z} = (x, \mu, \lambda)$. Let $\widetilde{\pi}$ be the canonical projection from P to $\widetilde{P}$. We then have following theorem:

Theorem 4.5.1: The Hamiltonian system $(P, \{ , \}_P, X_H)$ has $SO(3)$ symmetry and is Poisson reducible. The Poisson reduced system is given by

$$\dot{\widetilde{z}} = \widetilde{\Lambda}(\widetilde{z}) \nabla \widetilde{H}, \tag{4.5.4}$$

where $\widetilde{\Lambda}(\widetilde{z})$ is the reduced Poisson structure given by

$$\widetilde{\Lambda}(\widetilde{z}) = \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & \widehat{\lambda} \end{pmatrix} \tag{4.5.5}$$

98

for I being 2×2 identical matrix and 0's null matrices with suitable dimensions, and $\widetilde{H}$ the reduced Hamiltonian given by

$$\widetilde{H} \circ \widetilde{\pi} = H.$$

In addition, the Casimir functions for the Poisson reduced system are all real-valued smooth functions of $\|\lambda\|^2$ which is a first integral for the system, i.e., for some constant C_1,

$$\|\lambda\|^2 = C_1. \tag{4.5.6}$$

Proof: It is obvious that the Hamiltonian (4.5.2) is invariant under the action Φ^{T^*} since it does not have matrix A in its expression. This immediately implies the $SO(3)$-symmetry for the system. Moreover, the reduced Hamiltonian $\widetilde{H}$ on $T^*(\Re^2 \times SO(3))/SO(3)$ is simply

$$\widetilde{H}(x, \mu, \lambda) = H(x, A, \mu, A\widehat{\lambda}). \tag{4.5.7}$$

Let f_1 and f_2 be smooth functions on $\widetilde{P} = T^*(\Re^2 \times SO(3))/SO(3)$. Let F_1 and F_2 be lifed functions on $P = T^*(\Re^2 \times SO(3))$ such that

$$F_i(x, A, \mu, A\widehat{\lambda}) = f_i(x, \mu, \lambda), \quad i = 1, 2.$$

We need to find the expression of $\{f_1, f_2\}_{\widetilde{P}}$ such that

$$\{f_1, f_2\}_{\widetilde{P}}(x, \mu, \lambda) = \{F_1, F_2\}_P(x, A, \mu, A\widehat{\lambda}). \tag{4.5.8}$$

where $\{F_1, F_2\}_P$ is given in (4.5.1). As we have seen in the second proof of Theorem 4.4.5, a tangent vector Y on $T^*(\Re^2 \times SO(3))$ at $z = (x, A, \mu, A\widehat{\lambda})$ has the form

$$Y(z) = (v, A\widehat{\alpha}, w, A(\widehat{\alpha}\widehat{\lambda} + \widehat{\beta}))_z,$$

where $v, w \in \Re^2$ and $\alpha, \beta \in \Re^3$. An integral curve of Y at z can be written as

$$y(t) = (x + tv, Ae^{t\widehat{\alpha}}, \mu + tw, Ae^{t\widehat{\alpha}}(\widehat{\lambda} + t\widehat{\beta})).$$

Then, by the definition of F_i, we have

$$\begin{aligned}
dF_i \cdot Y(z) &= \frac{d}{dt}\Big|_{t=0} F_i(x + tv, Ae^{t\widehat{\alpha}}, \mu + tw, Ae^{t\widehat{\alpha}}(\widehat{\lambda} + t\widehat{\beta})) \\
&= \frac{d}{dt}\Big|_{t=0} f_i(x + tv, \mu + tw, \lambda + t\beta) \\
&= \frac{\partial f_i}{\partial x} \cdot v + \frac{\partial f_i}{\partial \mu} \cdot w + \frac{\partial f_i}{\partial \lambda} \cdot \beta.
\end{aligned} \tag{4.5.9}$$

99

On the other hand

$$dF_i \cdot Y(z) = \frac{\partial F_i}{\partial x} \cdot v + \langle D_A F_i, A\alpha \rangle + \frac{\partial F_i}{\partial \mu} \cdot w + \langle D_{A\widehat{\lambda}} F_i, A(\widehat{\alpha}\widehat{\lambda} + \beta) \rangle$$

$$= \frac{\partial F_i}{\partial x} \cdot v + \frac{\partial F_i}{\partial \mu} \cdot w + \langle D_A F_i - (D_{A\widehat{\lambda}} F_i)\widehat{\lambda}, A\widehat{\alpha} \rangle + \langle D_{A\widehat{\lambda}} F_i, A\widehat{\beta} \rangle.$$

(4.5.10)

Comparing (4.5.9) with (4.5.10), we get

$$\frac{\partial F_i}{\partial x} = \frac{\partial f_i}{\partial x}, \qquad \frac{\partial F_i}{\partial \mu} = \frac{\partial f_i}{\partial \mu},$$

(4.5.11a)

$$D_{A\widehat{\lambda}} F_i = A \widehat{\frac{\partial f_i}{\partial \lambda}}, \qquad D_A F_i = A \widehat{\frac{\partial f_i}{\partial \lambda}} \widehat{\lambda}.$$

(4.5.11b)

From (4.5.1), (4.5.8) and (4.5.11), we have

$$\{f_1, f_2\}_{\widetilde{P}}(\widetilde{z}) = \langle A \widehat{\frac{\partial f_1}{\partial \lambda}} \widehat{\lambda}, A \widehat{\frac{\partial f_2}{\partial \lambda}} \rangle - \langle A \widehat{\frac{\partial f_2}{\partial \lambda}} \widehat{\lambda}, A \widehat{\frac{\partial f_1}{\partial \lambda}} \rangle$$

$$+ \frac{\partial f_1}{\partial x} \cdot \frac{\partial f_2}{\partial \mu} - \frac{\partial f_2}{\partial x} \cdot \frac{\partial f_1}{\partial \mu}$$

$$= -(\frac{\partial f_1}{\partial \lambda} \times \frac{\partial f_2}{\partial \lambda}) \cdot \lambda + \frac{\partial f_1}{\partial x} \cdot \frac{\partial f_2}{\partial \mu} - \frac{\partial f_2}{\partial x} \cdot \frac{\partial f_1}{\partial \mu}$$

(4.5.12)

$$= \begin{pmatrix} \frac{\partial f_1}{\partial x}^T & \frac{\partial f_1}{\partial \mu}^T & \frac{\partial f_1}{\partial \lambda}^T \end{pmatrix} \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & \widehat{\lambda} \end{pmatrix} \begin{pmatrix} \frac{\partial f_2}{\partial x} \\ \frac{\partial f_2}{\partial \mu} \\ \frac{\partial f_2}{\partial \lambda} \end{pmatrix},$$

where I is the 2×2 identity matrix. Therefore, the matrix in (4.5.5) is reduced Poisson structure. The proof of the rest of this theorem can be easily carried out. ∎

Since we are dealing with a trivial principal bundle here with structure group as the symmetry group, after recognizing the coordinates for $\widetilde{P}$ one should be able to find the Poisson reduced system by eliminating the third equation of (4.4.10a) and determine the reduced Poisson structure from it. The first integral in (4.5.6) can be easily determined from the last equation of (4.4.10a).

Following Remark 4.4.6, it is clear that this reduction theorem does not depend on the forms of vectors Ω_1 and Ω_2, either. It can also be extended to a system on a principal bundle $(\Re^m \times SO(3), \Re^m, \pi, SO(3))$.

4.6 Approximation and Further Reduction

Up to now, we have made no simplifications or approximations of the optimal control problem. It is customary to make *ad hoc* approximations in the interest of

ensuring analytic integrability or numerical solvability. However, in the process one can easily destroy symmetries inherent to the problem. This is highly undesirable. On the other hand, in many physical systems, simplification and/or approximation may bring symmetries to the system. Our purpose in this section is to impose suitable assumptions, explore further symmetry and reduce the order of the system again.

As having done in Section 4.3, we assume that the distances of the point masses to 0-z axis, x_i, are very small in comparison with the distances of them to 0-xy plan, l, i.e.,

$$|x_i|/l \ll 1. \tag{4.6.1}$$

By doing so, we ignore the higher order terms with $(\frac{x_1}{l})^i(\frac{x_2}{l})^j$ for $i + j > 2$ in both numerators and denominators of $\Omega_1(x_1, x_2)$ and $\Omega_2(x_1, x_2)$. Then the approximate angular velocity (using the same notation) of the rigid body becomes

$$\Omega = \begin{pmatrix} 0 \\ a_1 \\ b_1 x_2 \end{pmatrix} \dot{x}_1 + \begin{pmatrix} a_2 \\ 0 \\ b_2 x_1 \end{pmatrix} \dot{x}_2 \tag{4.6.2}$$

where

$$\begin{aligned}
a_1 &= -\frac{1}{\Delta} \epsilon m I_z l (2\epsilon m l^2 + I_x) \\
b_1 &= \frac{1}{\Delta} \epsilon^2 m (2\epsilon m l^2 + I_x)(2\epsilon m l^2 - m l^2 + I_y) \\
a_2 &= -\frac{1}{\Delta} \epsilon m I_z l (2\epsilon m l^2 + I_y) \\
b_2 &= -\frac{1}{\Delta} \epsilon^2 m (2\epsilon m l^2 + I_y)(2\epsilon m l^2 - m l^2 + I_x)
\end{aligned} \tag{4.6.3}$$

for

$$\Delta = I_z(2\epsilon m l^2 + I_x)(2\epsilon m l^2 + I_y).$$

Remark 4.6.1: From (4.6.2), we see that if second point mass is fixed on 0-z axis, i.e. $x_2 = 0$, an infinitesimal linear motion of first point mass along 0-x axis will generate infinitesimal rotational motion of rigid body about 0-y axis only. However, if there is an small offest of second masses, i.e. $x_2 \neq 0$, an infinitesimal linear motion of first point mass along 0-x axis will generate infinitesimal rotating motion of rigid body about both 0-y axis and 0-z axis. Therefore, the above assumption does make physical sense. ∎

101

Inspired by the symmetric heavy top, we pose one more important assumption. We assume that the rigid body is symmetric about $0\text{-}z$ axis which implies

$$I_x = I_y. \tag{4.6.4}$$

Under this assumption Ω takes the form:

$$\Omega = \begin{pmatrix} 0 \\ a \\ bx_2 \end{pmatrix} \dot{x}_1 + \begin{pmatrix} a \\ 0 \\ -bx_1 \end{pmatrix} \dot{x}_2, \tag{4.6.5}$$

where $a = a_1$ and $b = b_1$. The approximation of $\widetilde{H}$ is, for $\widetilde{z} = (x, \mu, \lambda)$,

$$H_{12}(\widetilde{z}) = -\frac{1}{4}((b\lambda_3 x_2 + \mu_1 + a\lambda_2)^2 + (b\lambda_3 x_1 - \mu_2 - a\lambda_1)^2). \tag{4.6.6}$$

Applying H_{12} to (4.5.4), we get corresponding Poisson dynamics

$$\dot{\widetilde{z}} = \widetilde{\Lambda}(\widetilde{z}) \nabla H_{12}$$

or, explicitly,

$$\begin{cases}
\dot{x}_1 = -\dfrac{b\lambda_3 x_2 + \mu_1 + a\lambda_2}{2} \\[2mm]
\dot{x}_2 = \dfrac{b\lambda_3 x_1 - \mu_2 - a\lambda_1}{2} \\[2mm]
\dot{\mu}_1 = \dfrac{b^2\lambda_3^2 x_1 - b\lambda_3\mu_2 - ab\lambda_1\lambda_3}{2} \\[2mm]
\dot{\mu}_2 = \dfrac{b^2\lambda_3^2 x_2 + b\lambda_3\mu_1 + ab\lambda_2\lambda_3}{2} \\[2mm]
\dot{\lambda}_1 = -\dfrac{1}{2}(b^2\lambda_2\lambda_3 x_2^2 + (b\lambda_2\mu_1 - ab\lambda_3^2 + ab\lambda_2^2)x_2 + b^2\lambda_2\lambda_3 x_1^2 \\[1mm]
\qquad\quad + (-b\lambda_2\mu_2 - ab\lambda_1\lambda_2)x_1 - a\lambda_3\mu_1 - a^2\lambda_2\lambda_3) \\[2mm]
\dot{\lambda}_2 = \dfrac{1}{2}(b^2\lambda_1\lambda_3 x_2^2 + (b\lambda_1\mu_1 + ab\lambda_1\lambda_2)x_2 + b^2\lambda_1\lambda_3 x_1^2 \\[1mm]
\qquad\quad + (-b\lambda_1\mu_2 + ab\lambda_3^2 - ab\lambda_1^2)x_1 - a\lambda_3\mu_2 - a^2\lambda_1\lambda_3) \\[2mm]
\dot{\lambda}_3 = -\dfrac{ab\lambda_1\lambda_3 x_2 + ab\lambda_2\lambda_3 x_1 - a\lambda_2\mu_2 + a\lambda_1\mu_1}{2}
\end{cases} \tag{4.6.7}$$

Remark 4.6.2: A standard way to approximate Hamiltonian system is to keep some terms of Taylor expansion of the Hamiltonian up to a certain order. Our approximation above keeps 0th, 1st and part of 2nd order terms of Taylor expansion of $\widetilde{H}$. Its physical meaning is given in Remark 4.6.1. ∎

Next, we show that the above system admits a symmetry group and the order of the reduced Hamiltonian system can be finally brought down to 4. Consider a one-parameter group $G_\tau \simeq S^1$ with each element having the form:

$$g_\tau = BlockDiag(Rot(\tau), Rot(\tau), Rot^3(-\tau)), \qquad (4.6.8)$$

where

$$Rot(\tau) = \begin{pmatrix} cos(\tau) & sin(\tau) \\ -sin(\tau) & cos(\tau) \end{pmatrix} \quad \text{and} \quad Rot^3(\tau) = \begin{pmatrix} cos(\tau) & sin(\tau) & 0 \\ -sin(\tau) & cos(\tau) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Define an action Ψ of G_τ on $\widetilde{P} = T^*(\Re^2 \times SO(3))/SO(3) \simeq \Re^7$ by

$$\Psi : G_\tau \times \Re^7 \to \Re^7$$

$$(\tau, (x, \mu, \lambda)) \mapsto (Rot(\tau)x, Rot(\tau)\mu, Rot^3(-\tau)\lambda). \qquad (4.6.9)$$

We then have following striking fact.

Theorem 4.6.3: Following the rule of (4.6.9), the group G_τ acts on the Poisson manifold $\widetilde{P}$ canonically, i.e., for any $g_\tau \in G_\tau$,

$$\{f_1, f_2\}_{\widetilde{P}} \circ \Psi_{g_\tau} = \{f_1 \circ \Psi_{g_\tau}, f_2 \circ \Psi_{g_\tau}\}_{\widetilde{P}}, \qquad \forall f_i \in C^\infty(\widetilde{P}), \forall g_\tau \in G_\tau.$$

In addition, the approximate Hamiltonian H_{12} is G_τ-invariant, i.e.

$$H_{12}(\widetilde{z}) = H_{12}(\Psi_{g_\tau}(\widetilde{z})).$$

Proof: The first assertion is equivalent to

$$D\Psi_{g_\tau}(\widetilde{z})\widetilde{\Lambda}(\widetilde{z})D\Psi_{g_\tau}(\widetilde{z})^T = \widetilde{\Lambda}(\Psi_{g_\tau}(\widetilde{z})).$$

This can be shown by a straightforward calculation. So is the second assertion. ∎

Remark 4.6.4: From this theorem, one immediately concludes that G_τ is a symmetric group of the system (4.6.7). ∎

It is clear that by Theorem 4.6.3 the conditions of Noether's theorem introduced in Section 2.1 are satisfied and, consequently, the associated momentum map is an integral of the reduced system (4.6.7). From Section 2.1 we know that the momentum map $\mathbf{J} : \widetilde{P} \to \mathcal{G}_\tau^*$ (the dual of Lie algebra of G_τ) of this action is defined by

$$\langle \mathbf{J}(\widetilde{z}), \xi \rangle = J(\xi)(\widetilde{z}) \qquad (4.6.10)$$

103

for all $\xi \in \mathcal{G}_\tau$ and $\widetilde{z} \in \widetilde{P}$, where $J : \mathcal{G}_\tau \to C^\infty(\widetilde{P})$ is a linear map such that

$$X_{J(\xi)} = \xi_{\widetilde{P}}. \qquad (4.6.11)$$

Theorem 4.6.5: The constant momentum map for the system $(\widetilde{P}, \{\ ,\ \}_{\widetilde{P}})$ corresponding to the action Ψ of G_τ on $\widetilde{P}$ is

$$J(\xi)(\widetilde{z}) = -x_1\mu_2 + x_2\mu_1 + \lambda_3. \qquad (4.6.12)$$

Proof:

Recall that the Hamiltonian vector field $X_{H_{12}}$ on $\widetilde{P}$ is given by

$$X_{H_{12}}(f) = \{f, H_{12}\}_{\widetilde{P}} \qquad \forall f \in C^\infty(\widetilde{P}),$$

and from (4.6.11), we can determine the momentum map by

$$\{f, J(\xi)\}_{\widetilde{P}} = \xi_{\widetilde{P}}(f) \qquad \forall \xi \in \mathcal{G}_\tau.$$

From (4.6.8) we know that $\xi \in \mathcal{G}_\tau$ is of the form

$$\xi = \xi' BlockDiag\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right),$$

where ξ' is any constant in $\Re$ which will be chosen as 1 later. It is easy to show that the infinitesimal generator of the action (4.6.9) corresponding to ξ is given by

$$\xi_{\widetilde{P}}(\widetilde{z}) = \left.\frac{d}{d\tau}\right|_{\tau=0} \Psi(exp(\xi\tau), \widetilde{z}) = (x_2, -x_1, \mu_2, -\mu_1, -\lambda_2, \lambda_1, 0)^T,$$

where Ψ is defined in (4.6.9). Then, for any function f on $\widetilde{P}$,

$$\xi_{\widetilde{P}}(f)(\widetilde{z}) = x_2 \frac{\partial f}{\partial x_1} - x_1 \frac{\partial f}{\partial x_2} + \mu_2 \frac{\partial f}{\partial \mu_1} - \mu_1 \frac{\partial f}{\partial \mu_2} - \lambda_2 \frac{\partial f}{\partial \lambda_1} + \lambda_1 \frac{\partial f}{\partial \lambda_2} + 0 \frac{\partial f}{\partial \lambda_3}. \quad (4.6.13)$$

On the other hand, let $J(\xi)$ be a function on M, then

$$\begin{aligned}
\{f, J(\xi)\}(\widetilde{z}) = {}& df(\widetilde{z})^T \widetilde{\Lambda}(\widetilde{z}) dJ(\xi)(\widetilde{z}) \\
= {}& \frac{\partial J(\xi)}{\partial \mu_1} \frac{\partial f}{\partial x_1} + \frac{\partial J(\xi)}{\partial \mu_2} \frac{\partial f}{\partial x_2} - \frac{\partial J(\xi)}{\partial x_1} \frac{\partial f}{\partial \mu_1} - \frac{\partial J(\xi)}{\partial x_2} \frac{\partial f}{\partial \mu_2} \\
& + (\lambda_3 \frac{\partial J(\xi)}{\partial \lambda_2} - \lambda_2 \frac{\partial J(\xi)}{\partial \lambda_3}) \frac{\partial f}{\partial \lambda_1} + (-\lambda_3 \frac{\partial J(\xi)}{\partial \lambda_1} + \lambda_1 \frac{\partial J(\xi)}{\partial \lambda_3}) \frac{\partial f}{\partial \lambda_2} \\
& + (\lambda_2 \frac{\partial J(\xi)}{\partial \lambda_1} - \lambda_1 \frac{\partial J(\xi)}{\partial \lambda_2}) \frac{\partial f}{\partial \lambda_3}.
\end{aligned} \qquad (4.6.14)$$

Comparing (4.6.13) and (4.6.14), we have following PDE

$$\frac{\partial J(\xi)}{\partial \mu_1} = x_2$$

$$\frac{\partial J(\xi)}{\partial \mu_2} = -x_1$$

$$-\frac{\partial J(\xi)}{\partial x_1} = \mu_2$$

$$-\frac{\partial J(\xi)}{\partial x_2} = -\mu_1$$

$$\lambda_3 \frac{\partial J(\xi)}{\partial \lambda_2} - \lambda_2 \frac{\partial J(\xi)}{\partial \lambda_3} = -\lambda_2$$

$$-\lambda_3 \frac{\partial J(\xi)}{\partial \lambda_1} + \lambda_1 \frac{\partial J(\xi)}{\partial \lambda_3} = \lambda_1$$

$$\lambda_2 \frac{\partial J(\xi)}{\partial \lambda_1} - \lambda_1 \frac{\partial J(\xi)}{\partial \lambda_2} = 0.$$

One can check that function

$$J(\xi)(\widetilde{z}) = -x_1\mu_2 + x_2\mu_1 + \lambda_3$$

is a solution of the above PDE. Therefore, from Noether's theorem, this function is a constant along the motion of $X_{H_{12}}$, i.e.,

$$-x_1\mu_2 + x_2\mu_1 + \lambda_3 = C_2 \qquad (4.6.15)$$

for some constant C_2. ∎

Since the reduced system has G_τ-symmetry, by using standard Poisson reduction procedure again, we can drop the system (4.6.7) to quotient space $\bar{P} \overset{\triangle}{=} \widetilde{P}/G_\tau \simeq T^*(\Re^2 \times SO(3))/SO(3)/S^1$ with projection $\bar{\pi} : \widetilde{P} \to \bar{P}$. In the following, we will find induced Hamiltonian $\bar{H}_{12}$, induced Poisson structure $\bar{\Lambda}$ and reduced Hamiltonian vector field $X_{\bar{H}_{12}}$ on manifold $\bar{P}$.

First, consider a change of coordinates

$$\psi : \widetilde{P} \to \widetilde{P}$$

$$(4.6.16a)$$

$$\widetilde{z} = (x_1, x_2, \mu_1, \mu_2, \lambda_1, \lambda_2, \lambda_3) \mapsto \widetilde{z}' = (r_1, r_2, r_3, \theta_1, \theta_2, \theta_3, \lambda_3)$$

given by relations

$$\begin{cases} x_1 = r_1 cos(\theta_1), \\ x_2 = r_1 sin(\theta_1), \\ \mu_1 = r_2 cos(\theta_2), \\ \mu_2 = r_2 sin(\theta_2), \end{cases} \qquad \begin{cases} \lambda_1 = r_3 cos(\theta_3), \\ \lambda_2 = -r_3 sin(\theta_3), \\ \lambda_3 = \lambda_3. \end{cases} \qquad (4.6.16b)$$

With this set of new coordinates, the Hamiltonian H_{12} becomes

$$H'_{12}(\tilde{z}') = \frac{1}{4}(2ar_2r_3sin(\theta_3 - \theta_2) + 2ab\lambda_3r_1r_3cos(\theta_3 - \theta_1)$$
$$+ 2b\lambda_3r_1r_2sin(\theta_2 - \theta_1) - a^2r_3^2 - r_2^2 - b^2\lambda_3^2r_1^2).$$

(4.6.17)

And the corresponding dynamics $X_{H'_{12}}$ is given by

$$\begin{cases}
\dot{r}_1 = \dfrac{ar_3sin(\theta_3 - \theta_1) - r_2cos(\theta_2 - \theta_1)}{2} \\[2mm]
\dot{\theta}_1 = -\dfrac{ar_3cos(\theta_3 - \theta_1) + r_2sin(\theta_2 - \theta_1) - b\lambda_3r_1}{2r_1} \\[2mm]
\dot{r}_2 = -\dfrac{ab\lambda_3r_3cos(\theta_3 - \theta_2) - b^2\lambda_3^2r_1cos(\theta_2 - \theta_1)}{2} \\[2mm]
\dot{\theta}_2 = -\dfrac{1}{2r_2}(ab\lambda_3r_3sin(\theta_3 - \theta_2) + b^2\lambda_3^2r_1sin(\theta_2 - \theta_1) - b\lambda_3r_2) \\[2mm]
\dot{r}_3 = \dfrac{a\lambda_3r_2cos(\theta_3 - \theta_2) - ab\lambda_3^2r_1sin(\theta_3 - \theta_1)}{2} \\[2mm]
\dot{\theta}_3 = -\dfrac{1}{2r_3}(a\lambda_3r_2sin(\theta_3 - \theta_2) + (ab\lambda_3^2r_1 - abr_1r_3^2)cos(\theta_3 - \theta_1) \\[2mm]
\qquad\quad - br_1r_2r_3sin(\theta_2 - \theta_1) + (b^2\lambda_3r_1^2 - a^2\lambda_3)r_3) \\[2mm]
\dot{\lambda}_3 = -\dfrac{ar_2r_3cos(\theta_3 - \theta_2) - ab\lambda_3r_1r_3sin(\theta_3 - \theta_1)}{2}.
\end{cases}$$

(4.6.18)

Observing that the Hamiltonian H'_{12} and the right-hand-side of differential equation (4.8.18) depend on relative values of θ_1, θ_2 and θ_3 only, we can reduce the order of this system as follows. Let

$$\theta_{21} = \theta_2 - \theta_1$$

$$\theta_{32} = \theta_3 - \theta_2.$$

By using $\bar{z} = (r_1, r_2, r_3, \theta_{21}, \theta_{32}, \lambda_3)$ to parameterize $\bar{P}$, the induced Hamiltonian on $\bar{P}$ is given by

$$\bar{H}_{12}(\bar{z}) = \frac{1}{4}(2ab\lambda_3r_1r_3cos(\theta_{32} + \theta_{21}) + 2ar_2r_3sin(\theta_{32})$$
$$+ 2b\lambda_3r_1r_2sin(\theta_{21}) - a^2r_3^2 - r_2^2 - b^2\lambda_3^2r_1^2).$$

(4.6.19)

The corresponding induced dynamics $X_{\bar{H}_{12}}$ on $\bar{P}$ is given by

$$
\begin{cases}
\dot{r}_1 = \dfrac{ar_3 sin(\theta_{32} + \theta_{21}) - r_2 cos(\theta_{21})}{2} \\[2mm]
\dot{r}_2 = -\dfrac{ab\lambda_3 r_3 cos(\theta_{32}) - b^2\lambda_3^2 r_1 cos(\theta_{21})}{2} \\[2mm]
\dot{r}_3 = -\dfrac{ab\lambda_3^2 r_1 sin(\theta_{32} + \theta_{21}) - a\lambda_3 r_2 cos(\theta_{32})}{2} \\[2mm]
\dot{\theta}_{21} = \dfrac{1}{2r_1 r_2}(ar_2 r_3 cos(\theta_{32} + \theta_{21}) - ab\lambda_3 r_1 r_3 sin(\theta_{32}) \\[2mm]
\qquad + (r_2^2 - b^2\lambda_3^2 r_1^2)sin(\theta_{21})) \\[2mm]
\dot{\theta}_{32} = \dfrac{1}{2r_2 r_3}((abr_1 r_2 r_3^2 - ab\lambda_3^2 r_1 r_2)cos(\theta_{32} + \theta_{21}) \\[2mm]
\qquad + (ab\lambda_3 r_3^2 - a\lambda_3 r_2^2)sin(\theta_{32}) \\[2mm]
\qquad + (br_1 r_2^2 + b^2\lambda_3^2 r_1)r_3 sin(\theta_{21}) \\[2mm]
\qquad + ((a^2 - b)\lambda_3 - b^2\lambda_3 r_1^2)r_2 r_3) \\[2mm]
\dot{\lambda}_3 = \dfrac{ab\lambda_3 r_1 r_3 sin(\theta_{32} + \theta_{21}) - ar_2 r_3 cos(\theta_{32})}{2}
\end{cases}
\tag{4.6.20}
$$

Moreover, the first integrals in (4.5.6) and (4.6.15) now take form:

$$
r_3^2 + \lambda_3^2 = C_1
\tag{4.6.21}
$$

and

$$
r_1 r_2 sin(\theta_{21}) + \lambda_3 = C_2.
\tag{4.6.22}
$$

Therefore, as we claimed before, the final reduced system with the above two integrals is a four-dimensional Hamiltonian system.

Next, we show that the final reduced system (4.6.20) with (4.6.21) and (4.6.22) is also Poisson. In other words, one should be able to find a Poisson structure, $\bar{\Lambda}$, on $\bar{P}$ such that equations (4.6.20) can be expressed as

$$
\dot{\bar{z}} = \bar{\Lambda}(\bar{z})\nabla \bar{H}_{12}(\bar{z}).
$$

To this end, we first look at the Poisson tensor $\widetilde{\Lambda}$ under the new coordinates $\widetilde{z}' = (r_1, r_2, r_3, \theta_1, \theta_2, \theta_3, \lambda_3)$ in $\widetilde{P}$. Since $\widetilde{z}' = \psi(\widetilde{z})$ for $\widetilde{z} = (x_1, x_2, \mu_1, \mu_2, \lambda_1, \lambda_2, \lambda_3)$, the Poisson bracket under new coordinates will satisfy

$$
\{F_1, F_2\}(\widetilde{z}') = \{f_1, f_2\} \circ \psi^{-1}(\widetilde{z}'),
\tag{4.6.23}
$$

where $F_i = f_i \circ \psi^{-1}$ for $f_i \in C^\infty(\widetilde{P})$, $i = 1, 2$. The right-hand-side of (4.6.23) gives

$$\{f_1, f_2\} \circ \psi^{-1}(\widetilde{z}') = \frac{\partial f_1}{\partial \widetilde{z}}^T\Big|_{\widetilde{z} = \psi^{-1}(\widetilde{z}')} \Lambda(\psi^{-1}(\widetilde{z}')) \frac{\partial f_1}{\partial \widetilde{z}}\Big|_{\widetilde{z} = \psi^{-1}(\widetilde{z}')}.$$

It is easy to show that

$$\frac{\partial f_i}{\partial \widetilde{z}} = (\frac{\partial \psi}{\partial \widetilde{z}})^T\Big|_{\widetilde{z} = \psi^{-1}(\widetilde{z}')} \frac{\partial F_i}{\partial \widetilde{z}'}.$$

Therefore, the Poisson tensor under new coordinates is given by

$$\widetilde{\Lambda}'(\widetilde{z}') = (\frac{\partial \psi}{\partial \widetilde{z}})\Big|_{\widetilde{z} = \psi^{-1}(\widetilde{z}')} \widetilde{\Lambda}(\psi^{-1}(\widetilde{z}'))(\frac{\partial \psi}{\partial \widetilde{z}})^T\Big|_{\widetilde{z} = \psi^{-1}(\widetilde{z}')}.$$

From (4.6.16), one can show that

$$\frac{\partial \psi}{\partial \widetilde{z}}\Big|_{\widetilde{z} = \psi^{-1}(\widetilde{z}')} = \begin{pmatrix} cos(\theta_1) & sin(\theta_1) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & cos(\theta_2) & sin(\theta_2) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & cos(\theta_3) & -sin(\theta_3) & 0 \\ -\frac{sin(\theta_1)}{r_1} & \frac{cos(\theta_1)}{r_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{sin(\theta_2)}{r_2} & \frac{cos(\theta_2)}{r_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{sin(\theta_3)}{r_3} & -\frac{cos(\theta_3)}{r_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\widetilde{\Lambda}'(\widetilde{z}') = \begin{pmatrix} 0 & cos(\theta_2 - \theta_1) & 0 & 0 & -\frac{sin(\theta_2 - \theta_1)}{r_2} & 0 & 0 \\ -cos(\theta_2 - \theta_1) & 0 & 0 & -\frac{sin(\theta_2 - \theta_1)}{r_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\lambda_3}{r_3} & 0 \\ 0 & \frac{sin(\theta_2 - \theta_1)}{r_1} & 0 & 0 & \frac{cos(\theta_2 - \theta_1)}{r_1 r_2} & 0 & 0 \\ \frac{sin(\theta_2 - \theta_1)}{r_2} & 0 & 0 & -\frac{cos(\theta_2 - \theta_1)}{r_1 r_2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{\lambda_3}{r_3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \quad (4.6.24)$$

which depends on $\theta_2 - \theta_1$ only.

We now ready to determine the Poisson tensor on $\bar{P}$. Let $\bar{f}_1$ and $\bar{f}_2$ be smooth functions on $\bar{P}$ which is parameterized by $\bar{z} = (r_1, r_2, r_3, \theta_{21}, \theta_{32}, \lambda_3)$. Let $\widetilde{f}_1$ and $\widetilde{f}_2$ be lifed functions on $\widetilde{P}$ such that at lifted point $\bar{\pi}^{-1}(\bar{z}) = \widetilde{z} = (r_1, r_2, r_3, \theta_1, \theta_2, \theta_3, \lambda_3)$ in $\widetilde{P}$

$$\widetilde{f}_i(\widetilde{z}) = \bar{f}_i(\bar{z}), \qquad i = 1, 2.$$

The Poisson bracket of $\widetilde{f}_1$ and $\widetilde{f}_2$ at any point $\widetilde{z}$ in $\widetilde{P}$ is given by Poisson structure shown in (4.6.23). We need to find the expression of $\{\bar{f}_1, \bar{f}_2\}_{\widetilde{P}/G_\tau}$ such that

$$\{\bar{f}_1, \bar{f}_2\}_{\widetilde{P}/G_\tau} \circ \bar{\pi} = \{\widetilde{f}_1, \widetilde{f}_2\}_{\widetilde{P}}. \quad (4.6.25)$$

108

It is easy to show that

$$
\begin{pmatrix} \frac{\partial \widetilde{f}}{\partial r_1} \\ \frac{\partial \widetilde{f}}{\partial r_2} \\ \frac{\partial \widetilde{f}}{\partial r_3} \\ \frac{\partial \widetilde{f}}{\partial \theta_1} \\ \frac{\partial \widetilde{f}}{\partial \theta_2} \\ \frac{\partial \widetilde{f}}{\partial \theta_3} \\ \frac{\partial \widetilde{f}}{\partial \lambda_3} \end{pmatrix} (\widetilde{z}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\partial \bar{f}}{\partial r_1} \\ \frac{\partial \bar{f}}{\partial r_2} \\ \frac{\partial \bar{f}}{\partial r_3} \\ \frac{\partial \bar{f}}{\partial \theta_{21}} \\ \frac{\partial \bar{f}}{\partial \theta_{32}} \\ \frac{\partial \bar{f}}{\partial \lambda_3} \end{pmatrix} (\bar{z})
$$

or shortly,

$$
\frac{\partial \widetilde{f}}{\partial \widetilde{z}} = \Xi \frac{\partial \bar{f}}{\partial \bar{z}}.
$$

Then, immediately, from (4.6.23) and (4.6.24), we have

$$
\{\bar{f}_1, \bar{f}_2\}_{\widetilde{P}/G_\tau}(\bar{z}) = \frac{\partial \bar{f}_1}{\partial \bar{z}}^T \bar{\Lambda}(\bar{z}) \frac{\partial \bar{f}_2}{\partial \bar{z}},
$$

where

$$
\begin{aligned}
\bar{\Lambda}(\bar{z}) &= \Xi^T (\bar{\Lambda} \circ \bar{\pi}^{-1})(\bar{z}) \Xi \\
&= \begin{pmatrix} 0 & cos(\theta_{21}) & 0 & -\frac{sin(\theta_{21})}{r_2} & \frac{sin(\theta_{21})}{r_2} & 0 \\ -cos(\theta_{21}) & 0 & 0 & \frac{sin(\theta_{21})}{r_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\lambda_3}{r_3} \\ \frac{sin(\theta_{21})}{r_2} & -\frac{sin(\theta_{21})}{r_1} & 0 & 0 & \frac{cos(\theta_{21})}{r_1 r_2} & 0 \\ -\frac{sin(\theta_{21})}{r_2} & 0 & 0 & -\frac{cos(\theta_{21})}{r_1 r_2} & 0 & 0 \\ 0 & 0 & -\frac{\lambda_3}{r_3} & 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Therefore, (4.6.20) can be expressed by.

$$
\dot{\bar{z}} = \bar{\Lambda}(\bar{z}) \nabla \bar{H}_{12}(\bar{z})
$$

for $\bar{z} \in \bar{P}$.

Appendix 4.A:

This appendix gives complete expression of $\Omega_1(x_1, x_2)$ and $\Omega_2(x_1, x_2)$ in (4.4.1) with $\psi_1 = 0$ and $\psi_2 = \pi/2$.

$$\Omega_1 = \frac{1}{det(I_{lock})} \begin{pmatrix} \omega_{11} \\ \omega_{12} \\ \omega_{13} \end{pmatrix} \qquad \Omega_2 = \frac{1}{det(I_{lock})} \begin{pmatrix} \omega_{21} \\ \omega_{22} \\ \omega_{23} \end{pmatrix}$$

where

$$
\begin{aligned}
det(I_{lock}) =& (\epsilon - 1)\epsilon^3(2\epsilon - 1)m^3 x_1^2 x_2^4 \\
&+ (\epsilon - 1)\epsilon^2 m^2 (2\epsilon^2 mq^2 - \epsilon mq^2 + \epsilon I_y - I_y)x_2^4 \\
&+ (\epsilon - 1)\epsilon^3(2\epsilon - 1)m^3 x_1^4 x_2^2 \\
&+ \epsilon^2 m^2 (4\epsilon^3 mq^2 - 10\epsilon^2 mq^2 + 4\epsilon mq^2 - 2\epsilon I_z + I_z \\
&+ \epsilon^2 I_y - 2\epsilon I_y + I_y + \epsilon^2 I_x - 2\epsilon I_x + I_x)x_1^2 x_2^2 \\
&+ \epsilon m (4\epsilon^3 m^2 q^4 - 2\epsilon^2 m^2 q^4 + 2\epsilon^2 I_z mq^2 - 2\epsilon I_z mq^2 \\
&+ 2\epsilon^2 I_y mq^2 - 2\epsilon I_y mq^2 + 2\epsilon^2 I_x mq^2 \\
&- \epsilon I_x mq^2 + \epsilon I_y I_z - I_y I_z + \epsilon I_x I_y - I_x I_y)x_2^2 \\
&+ (\epsilon - 1)\epsilon^2 m^2 (2\epsilon^2 mq^2 - \epsilon mq^2 + \epsilon I_x - I_x)x_1^4 \\
&+ \epsilon m (4\epsilon^3 m^2 q^4 - 2\epsilon^2 m^2 q^4 + 2\epsilon^2 I_z mq^2 - 2\epsilon I_z mq^2 \\
&+ 2\epsilon^2 I_y mq^2 - \epsilon I_y mq^2 + 2\epsilon^2 I_x mq^2 \\
&- 2\epsilon I_x mq^2 + \epsilon I_x I_z - I_x I_z + \epsilon I_x I_y - I_x I_y)x_1^2 \\
&+ I_z(2\epsilon mq^2 + I_x)(2\epsilon mq^2 + I_y) \\
\omega_{11} =& - \epsilon^4(2\epsilon - 1)m^3 qx_1 x_2^3 \\
&- \epsilon^3 m^2 q(2\epsilon mq^2 - mq^2 - I_z + I_y)x_1 x_2 \\
\omega_{12} =& - (\epsilon - 1)\epsilon^3(2\epsilon - 1)m^3 qx_2^4 + \epsilon^3(2\epsilon - 1)m^2 qx_1^2 x_2^2 \\
&+ \epsilon^2 m^2 q(4\epsilon^2 mq^2 - 2\epsilon m^2 q^2 + \epsilon I_z - I_z + 2\epsilon I_x - I_x)x_2^2 \\
&+ \epsilon^2 mq(2\epsilon^2 mq^2 - \epsilon mq^2 + \epsilon I_x - I_x)x_1^2 \\
&- \epsilon m I_z q(2\epsilon mq^2 + I_x) \\
\omega_{13} =& - \epsilon^4(2\epsilon - 1)m^3 x_1^2 x_2^3
\end{aligned}
$$

$$- (\epsilon - 1)\epsilon^3 m^2 (2\epsilon m q^2 - m q^2 + I_y)x_2^3$$

$$- \epsilon^3 m^2 (2\epsilon^2 m q^2 - \epsilon m q^2 + \epsilon I_x - I_x)x_1^2 x_2$$

$$+ \epsilon^2 m (2\epsilon m q^2 + I_x)(2\epsilon m q^2 - m q^2 + I_y)x_2$$

$$\omega_{21} = \epsilon^3 (2\epsilon - 1)m^3 q x_1^2 x_2^2$$

$$+ \epsilon^2 m^2 q (2\epsilon^2 m q^2 - \epsilon m q^2 + \epsilon I_y - I_y)x_2^2$$

$$- (\epsilon - 1)\epsilon^3 (2\epsilon - 1)m^3 q x_1^4$$

$$+ \epsilon^2 m^2 q (4\epsilon^2 m q^2 - 2\epsilon m q^2 + \epsilon I_z - I_z + 2\epsilon I_y - I_y)x_1^2$$

$$- \epsilon m I_z q (2\epsilon m q^2 + I_y)$$

$$\omega_{22} = - \epsilon^4 (2\epsilon - 1)m^3 q x_1^3 x_2$$

$$- \epsilon^3 m^2 q (2\epsilon m q^2 - m q^2 - I_z + I_x)x_1 x_2$$

$$\omega_{23} = \epsilon^4 (2\epsilon - 1)m^3 x_1^3 x_2^2$$

$$+ \epsilon^3 m^2 (2\epsilon^2 m q^2 - \epsilon m q^2 + \epsilon I_y - I_y)x_1 x_2^2$$

$$+ (\epsilon - 1)\epsilon^3 m^2 (2\epsilon m q^2 - m q^2 + I_x)x_1^3$$

$$- \epsilon^2 m (2\epsilon m q^2 + I_y)(2\epsilon m q^2 - m q^2 + I_x)x_1.$$

CHAPTER V

CHAPLYGIN DYNAMICS AND LAGRANGIAN REDUCTION

In this chapter, based on the intrinsic form of *Lagrange-d'Alembert principle* formulated in Chapter II, we study constrained Lagrangian dynamics with symmetry on principal fiber bundles with connections. We show that, under certain constraints which are of the nonholonomic variety, there is a family of splittings of the Lagrangian velocity phase space. Each splitting gives a Lagrangian dynamics on the horizontal distribution given by a connection. If the nonholonomic constraint is given by the horizontal distribution of the connection itself, our result leads to Koiller's formula for non-Abelian Chaplygin systems [20]. If the mechanical connection is used and the exterior force leaves the momentum map invariant, our result leads to (non-Abelian) Lagrangian reduction due to Marsden and Scheurle [27].

In Section 5.1, as a motivation, we first consider a simple class of systems, namely, Lagrangian systems with Abelian symmetry and affine constraints. An important observation from studying such systems is that, when a constrained system possesses symmetry, the dynamics of the system can be described by a reduced system of lower dimension and without constraints. In Sections 5.2 and 5.3, we show that for a system

112

with non-Abelian symmetry such a reduction can also be obtained by going through a two-step procedure. In Section 5.4, we demonstrate Lagrangian reduction by using the formulas generated in Sections 5.2 and 5.3. The first section of Chapter IV provides some of the mathematical background (principal bundle, connection, etc.) for this chapter.

5.1 Abelian Chaplygin Systems with Affine Constraints

Let $Q = \Re^n = \Re^m \times \Re^{n-m}$ be the n-dimensional configuration space of the system under study, and $q = (q_1, q_2)$ a point in Q, for $q_1 \in \Re^m$ and $q_2 \in \Re^{n-m}$. Let function $L : TQ \to \Re$ be the Lagrangian of the system with the property:

$$L(q, v) = L((q_1, 0), (v_1, v_2)) \triangleq L(q_1, v_1, v_2), \qquad (5.1.1)$$

where $v = (v_1, v_2) \in T_q Q$, for $v_1 \in T_{q_1}\Re^m$ and $v_2 \in T_{q_2}\Re^{n-m}$. The constraint on the system is given by the zero level set of m linearly independent functions $f : TQ \to \Re^m$ which is of the form

$$f(q, v) = B(q_1)v_1 - v_2 + b(q_1), \qquad (5.1.2)$$

where $B : \Re^m \to \mathrm{L}(\Re^m, \Re^{n-m})$ and $b : \Re^m \to \Re^{n-m}$ are smooth mappings. In addition, the exterior force or control is given by a mapping $F : TQ \to T^*Q$ which is of the form

$$F(q, v) = (F_1(q_1, v_1, v_2), F_2(q_1, v_1, v_2)), \qquad (5.1.3)$$

for $F_1(q_1, v_1, v_2) \in T^*_{q_1}\Re^m$ and $F_2(q_1, v_1, v_2) \in T^*_{q_2}\Re^{n-m}$. Here, as well as in the rest of this chapter, defining the exterior force on TQ is equivalent to specifying a feedback control.

A general method to determine a motion of the system in the above setting is to solve the $(2n + m)$-dimensional differential equations given in (2.2.33) or (2.2.34). However, by observing the properties of the Lagrangian, the constraints and the exterior force given in (5.1.1)-(5.1.3), one can show that the problem can be reduced to solving a lower dimensional system of differential equations without constraints. Classically, a system as defined above is called a *Chaplygin system* provided that $b(q_1) \equiv 0$ [32]. It has been shown that for such systems, the dynamic equations can be reduced

to unconstrained equations on the space $\Re^m$ parametrized by q_1 with a modified Lagrangian *and* a modified exterior force. For the proof of this assertion and examples of physical systems, see [32]. From a geometric point of view, it is clear that the conditions for a system to be a Chaplygin system is that the Lagrangian, the exterior force and the constraints are invariant under the action of the Abelian group $\Re^{n-m}$, namely,

$$\Phi : \Re^{n-m} \times Q \to Q$$
$$(z, (q_1, q_2)) \mapsto (q_1, q_2 + z). \tag{5.1.4}$$

Because of this, we refer to a classical Chaplygin system as an *Abelian Chaplygin system with linear constraints* and the system given by (5.1.1)-(5.1.3) as an *Abelian Chaplygin system with affine constraints*. Next, we show how to derive reduced dynamic equations on $T\Re^m$ parametrized by (q_1, v_1) from the original system with affine constraints.

Consider an Abelian Chaplygin system with affine constraints given by (5.1.1)-(5.1.3). From *Lagrange-d'Alembert principle* formulated in (2.2.33), the dynamics is given by

$$\left(\frac{d}{dt} D_v L(q, v) - D_q L(q, v)\right) \cdot u = F(q, v) \cdot u \tag{5.1.5a}$$

with $(q, v) = ((q_1, q_2), (v_1, v_2)) \in TQ$ satisfying

$$f(q, v) = B(q_1)v_1 - v_2 + b(q_1) = 0 \tag{5.1.5b}$$

and $u = (u_1, u_2) \in T_q Q$ satisfying

$$D_v f(q, v) \cdot u = B(q_1)u_1 - u_2 = 0. \tag{5.1.5c}$$

Using (5.1.5c) and (5.1.1), Equation (5.1.5a) can be re-written as

$$\left(\frac{d}{dt} D_{v_1} L(q, v) - D_{q_1} L(q, v)\right) \cdot u_1 + \frac{d}{dt}(D_{v_2} L(q, v)) \cdot B(q_1)u_1$$
$$= \widetilde{F}(q_1, v_1) \cdot u_1, \tag{5.1.6}$$

where

$$\widetilde{F}(q_1, v_1) \stackrel{\triangle}{=} F_1(q_1, v_1, v_2)|_{v_2} + B(q_1)^* F_2(q_1, v_1, v_2)|_{v_2} \tag{5.1.7}$$

and $(\cdot)|_{v_2} \stackrel{\triangle}{=} (\cdot)|_{v_2 = B(q_1)v_1 + b(q_1)}$. Let

$$\widetilde{L}(q_1, v_1) \stackrel{\triangle}{=} L(q_1, v_1, B(q_1)v_1 + b(q_1)). \tag{5.1.8}$$

114

Then, it is easy to check the following equalities

$$D_{v_1}L(q_1, v_1, v_2) \cdot u_1 = D_{v_1}\widetilde{L}(q_1, v_1) \cdot u_1 - D_{v_2}L(q_1, v_1, v_2)|_{v_2} \cdot (B(q_1)u_1)$$

$$D_{q_1}L(q_1, v_1, v_2) \cdot u_1 = D_{q_1}\widetilde{L}(q_1, v_1) \cdot u_1 - D_{v_2}L(q_1, v_1, v_2)|_{v_2} \cdot ((D_{q_1}B(q_1)u_1)v_1)$$

$$- D_{v_2}L(q_1, v_1, v_2)|_{v_2} \cdot (D_{q_1}b(q_1)u_1),$$

where $D_{q_1}B(q_1) : \Re^m \to L(\Re^m, \Re^{n-m})$ and $D_{q_1}b(q_1) : \Re^m \to \Re^{n-m}$ are the Fréchet derivatives of $B(q_1)$ and $b(q_1)$, respectively. Substituting from the above two equations in (5.1.6), one gets

$$(\frac{d}{dt}D_{v_1}\widetilde{L}(q_1, v_1) - D_{q_1}\widetilde{L}(q_1, v_1)) \cdot u_1$$

$$= D_{v_2}L|_{v_2} \cdot [(D_{q_1}B(q_1)v_1)u_1 - (D_{q_1}B(q_1)u_1)v_1]$$

$$- D_{v_2}L|_{v_2} \cdot D_{q_1}b(q_1)u_1 + \widetilde{F}(q_1, v_1) \cdot u_1$$

$$= \langle D_{v_2}L|_{v_2}, dB(q_1)(v_1, u_1)\rangle$$

$$- \langle D_{v_2}L|_{v_2}, D_{q_1}b(q_1)u_1\rangle + \widetilde{F}(q_1, v_1) \cdot u_1,$$

where $dB(q_1)(v_1, u_1) \triangleq (D_{q_1}B(q_1)v_1)u_1 - (D_{q_1}B(q_1)u_1)v_1 \in \Re^{n-m}$. Writing things in terms of components, we have, for $B(q_1) = (B^{ij}, i = 1, \cdots, n-m; j = 1, \cdots, m)$, $v_1 = (v_1^k, k = 1, \cdots, m)$ and $u_1 = (u_1^l, l = 1, \cdots, m)$,

$$(dB(q_1)(v_1, u_1))^i = \sum_{k=1}^m \sum_{l=1}^m \left(\frac{\partial B^{ik}}{\partial q_1^l} - \frac{\partial B^{il}}{\partial q_1^k}\right) v_1^k u_1^l$$

for $i = 1, \cdots, n-m$.

We summarize the above result in the following theorem.

Theorem 5.1.1: If a curve $\{q(t) = (q_1(t), q_2(t)), t \geq 0\}$ in $Q = \Re^m \times \Re^{n-m}$ is a motion of the constrained Lagrangian system with the Lagrangian L, the constraints and the exterior force given in (5.1.1)-(5.1.3), respectively, then the curve $\{q_1(t), t \geq 0\}$ is the motion of the unconstrained Lagrangian system in $\Re^m$ with Lagrangian $\widetilde{L}$ given by (5.1.8) and the exterior forces given by $\langle D_{v_2}L|_{v_2}, dB(q_1)(v_1, \cdot)\rangle$, $-(D_{q_1}b(q_1))^* D_{v_2}L|_{v_2}$ and $\widetilde{F}(q_1, v_1)$ as in (5.1.7), i.e., it satisfies the reduced dynamics

$$\frac{d}{dt}D_{v_1}\widetilde{L}(q_1, v_1) \cdot u_1 - D_{q_1}\widetilde{L}(q_1, v_1) \cdot u_1 = \widetilde{F}(q_1, v_1) \cdot u_1 - \langle D_{v_2}L|_{v_2}, D_{q_1}b(q_1)u_1\rangle$$

$$+ \langle D_{v_2}L|_{v_2}, dB(q_1)(v_1, u_1)\rangle$$

(5.1.9)

for any $u_1 \in \Re^m$. ∎

115

Remarks 5.1.2:

(1) If $\{q_1(t), t \geq 0\}$ is known, the curve $\{q_2(t), t \geq 0\}$ can be determined uniquely from constraint equations by quadrature, i.e.,

$$q_2(t) = \int_0^t B(q_1(\tau))\dot{q}_1(\tau) + b(q_1(\tau))d\tau.$$

It is obvious that the curve $\{(q_1(t), q_2(t)), t \geq 0\}$ satisfies the original differential equation (5.1.5). Therefore, the condition (5.1.9) in Theorem 5.1.1 is also necessary.

(2) The results of Theorem 5.1.1 are applicable to a system with Lagrangian being not necessarily of the form, kinetic energy minus potential energy.

(3) If the configuration space Q in Theorem 5.1.1 is replaced by $B \times G$ where B is an m-dimensional manifold and G is an $(n-m)$-dimensional Abelian symmetry group of the system (i.e., $\Re^m$ is replaced by B and $\Re^{n-m}$ is replaced by G), then it is easy to see that Theorem 5.1.1 is still valid and Equation (5.1.9) is just the local form of the dynamics on B.

(4) An important observation here is that, as in the case of a Chaplygin system with *linear* constraints, when the system satisfies conditions (5.1.1)-(5.1.3), or *symmetry*, one can also simplify the problem of solving constrained dynamic equations given in (5.1.5) to the problem of solving the unconstrained dynamic equations given in (5.1.9) of lower dimension. As we shall see in the following sections, for systems with non-Abelian symmetry, such a simplification is also possible, but in two steps.

∎

Technically, the above derivation looks very elementary and quite the same as the one for a system with linear constraints. However, by applying this result to a system with a constant momentum map or a conservation law, we will discover the reduced dynamics originally derived by Routh [3,38], who obtained such equations by defining a new function, called the *Routhian*, through application of the so-called cyclic coordinates. It is our re-formulation of Routh's problem below via the nonholonomic constraints that motivated us to study more general systems, namely, non-Abelian Lagrangian reduction, as worked out in the later sections.

116

Consider a simple mechanical system with (Abelian) symmetry,

$$(Q = \Re^n, K, V, G = \Re^{n-m}). \tag{5.1.10}$$

Here, the Riemannian metric is written as

$$K(q)(v, w) = v^T M(q_1) w = (v_1{}^T, v_2{}^T) \begin{pmatrix} M_{11}(q_1) & M_{12}(q_1) \\ M_{12}(q_1)^T & M_{22}(q_1) \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

where $v_1, w_1 \in \Re^m$ and $v_2, w_2 \in \Re^{n-m}$, $M_{ij}(q_1)$ are the submatrices of the symmetric matrix M of suitable dimensions, and the potential function satisfies $V = V(q_1)$. The action of G on Q is given by (5.1.4). Then, the Lagrangian of the system is

$$\begin{aligned} L(q_1, v_1, v_2) =& \frac{1}{2} v^T M(q_1) v - V(q_1) \\ =& \frac{1}{2} v_1{}^T M_{11}(q_1) v_1 + v_1{}^T M_{12}(q_1) v_2 + \frac{1}{2} v_2{}^T M_{22}(q_1) v_2 - V(q_1). \end{aligned} \tag{5.1.11}$$

The Lie algebra of $G = \Re^{n-m}$ is $\mathcal{G} = \Re^{n-m}$. It is easy to check that, if $\xi \in \mathcal{G}$, the infinitesimal generator of the action Φ in (5.1.4) with respect to ξ is simply $\xi_Q(q) = (\mathbf{0}, \xi)$, where $\mathbf{0}$ is the null vector in $\Re^m$. Then, by (2.1.27), the momentum map with respect to the tangent action of Φ is

$$\mathbf{J}(q, v) = M_{12}^T(q_1) v_1 + M_{22}(q_1) v_2. \tag{5.1.12}$$

Assume that the exterior force or control leaves $\mu = \mathbf{J}(q, v)$ invariant. Since μ is fixed, the problem now is to determine the dynamics under the constraints given by the constant momentum map. Note that the Lagrangian given in (5.1.11) is G-invariant and $\mathbf{J}(q, v) = \mu$ can be re-arranged as

$$\begin{aligned} v_2 =& -M_{22}^{-1}(q_1) M_{12}^T(q_1) v_1 + M_{22}^{-1}(q_1) \mu \\ =& B(q_1) v_1 + b(q_1), \end{aligned} \tag{5.1.13}$$

where $B(q_1) \overset{\triangle}{=} -M_{22}^{-1}(q_1) M_{12}^T(q_1)$ and $b(q_1) \overset{\triangle}{=} M_{22}^{-1}(q_1) \mu$, which is also G-invariant. Therefore, we have an Abelian Chaplygin system with affine constraints. From the definition of $\widetilde{L}$ in (5.1.8), after a simple calculation, we have

$$\begin{aligned} \widetilde{L}(q_1, v_1) =& \frac{1}{2} v_1{}^T (M_{11} - M_{12} M_{22}^{-1} M_{12}^T) v_1 + \frac{1}{2} \mu^T M_{22}^{-1} \mu - V(q_1) \\ =& \frac{1}{2} v_1{}^T \widetilde{M} v_1 + \frac{1}{2} \mu^T M_{22}^{-1} \mu - V(q_1), \end{aligned}$$

where $\widetilde{M} \triangleq M_{11} - M_{12}M_{22}^{-1}M_{12}^T$,

$$D_{v_2}L(q_1, v_1, v_2)|_{v_2} = \mu,$$

and

$$\langle D_{v_2}L|_{v_2}, \; D_{q_1}b(q_1)u_1 \rangle = \langle \mu, \; (D_{q_1}M_{22}^{-1}(q_1)u_1)\mu \rangle = D_{q_1}(\mu^T M_{22}^{-1}(q_1)\mu) \cdot u_1.$$

From the above calculations, we see that if we define a function, referred to as the *reduced Lagrangian*, as

$$\widetilde{L}_\mu(q_1, v_1) \triangleq \frac{1}{2}v_1^T \widetilde{M}(q_1)v_1 - (\frac{1}{2}\mu^T M_{22}^{-1}(q_1)\mu + V(q_1)), \qquad (5.1.14)$$

equation (5.1.9) specializes to

$$\left(\frac{d}{dt}D_{v_1}\widetilde{L}_\mu(q_1, v_1) - D_{q_1}\widetilde{L}_\mu(q_1, v_1)\right) \cdot u_1 = \langle \mu, \; dB(q_1)(v_1, u_1) \rangle + \widetilde{F}(q_1, v_1) \cdot u_1. \quad (5.1.15)$$

Since the infinitesimal generator of the action in (5.1.4) corresponding to $\xi \in \mathcal{G}$ is $\xi_Q(q) = (0, \xi)$, the submatrix $\mathbb{I}(q_1) \triangleq M_{22}(q_1)$ is the locked inertia tensor (cf. Subsection 4.1.2). Then,

$$V_\mu \triangleq \frac{1}{2}\mu^T \mathbb{I}^{-1}(q_1)\mu + V(q_1)$$

is known as the *amended potential* (cf. [41]). In addition, from a geometric point of view, we claim that $-dB(q_1)$ is, in fact, the curvature form, Ω, of the mechanical connection on the principal bundle $(\Re^n, \Re^m, \Re^{n-m})$ and the first term in (5.1.15) is known as the μ-component of the curvature form evaluated at (u_1, v_1), i.e.,

$$\Omega_\mu(q_1)(u_1, v_1) \triangleq \langle \mu, \Omega(q_1)(u_1, v_1) \rangle = \langle \mu, \; -dB(q_1)(u_1, v_1) \rangle. \qquad (5.1.16)$$

Recall that the connection form $\omega(q)$ of the mechanical connection on Q is given as

$$\langle \omega(q), v \rangle = \mathbb{I}(q)^{-1}\mathbf{J}(q, v),$$

where $\mathbb{I}$ is locked inertia tensor and $\mathbf{J}$ is the momentum map. From the above discussion and (5.1.12) we have

$$\langle \omega(q), v \rangle = M_{22}^{-1}(q_1)M_{12}^T(q_1)v_1 + v_2.$$

118

This implies that the connection form is

$$\omega(q_1) = M_{22}^{-1}(q_1)M_{12}^T(q_1)dq_1 + dq_2$$

$$= -B(q_1)dq_1 + dq_2.$$

Since the symmetry group here is Abelian, by the definition of curvature form (cf. Subsection 4.1.1), the above equation justifies our claim.

In summary, we state the result as a theorem.

Theorem 5.1.3: If a curve $\{q(t) = (q_1(t), q_2(t)), t \geq 0\}$ in $Q = \Re^m \times \Re^{n-m}$ is a motion of the simple mechanical system with symmetry given in (5.1.10) together with exterior force (5.1.3) which leaves the momentum map given in (5.1.12) invariant, i.e., $\mathbf{J}(q, \dot{q}) = \mu$, then the curve $\{q_1(t), t \geq 0\}$ in $\Re^m$ satisfies Euler-Lagrange equation

$$\left(\frac{d}{dt}D_{v_1}\widetilde{L}_\mu(q_1, v_1) - D_{q_1}\widetilde{L}_\mu(q_1, v_1)\right) \cdot u_1 = \Omega_\mu(q_1)(u_1, v_1) + \widetilde{F}(q_1, v_1) \cdot u_1, \quad (5.1.17)$$

where Ω_μ is given in (5.1.16) and $\widetilde{F}(q_1, v_1)$ given in (5.1.7) for B therein given in (5.1.13). $\blacksquare$

Remarks 5.1.4:

(1) If Equation (5.1.15) or (5.1.17) is written in components, one gets the reduced dynamic equations due to Routh [38].

(2) Theorem 5.1.3 can be extended to a system on any n-dimensional smooth manifold Q, and $(n-m)$-dimensional Abelian symmetry group G. With this extension, Equation (5.1.17) gives the dynamics or special vector fields (cf. Chapter II), in local coordinates, on $T(Q/G)$. Then, the Theorem (5.1.3) is known as Lagrangian reduction for the system admitting Abelian symmetry, which is also shown in [3]. A natural question that follows is to work out the reduction theorem for the systems with non-Abelian symmetry. We will answer this question in Section 4. $\blacksquare$

5.2 Dynamics on Horizontal Distribution

Consider a simple mechanical system with symmetry given by a four-tuple

$$(Q, K, V, G), \quad (5.2.1)$$

where Q is the n-dimensional configuration space; G is a Lie group of dimension p acting on Q on the left freely and properly. This action is denoted by Φ given in (2.1.1); K is a Riemannian metric and G acts on Q by isometries; V is a G-invariant potential function (cf. Subsection 2.1.4). The Lagrangian of this system is given by

$$L(q, v_q) = \frac{1}{2} K(q)(v_q, v_q) - V(q). \tag{5.2.2}$$

In addition, we let the configuration space Q be a principal G'-bundle:

$$\wp = (Q, B, \pi, G'), \tag{5.2.3}$$

where G' is a p'-dimensional closed subgroup of G, $B = Q/G'$ is the $m(= n - p')$ dimensional base space and $\pi : Q \rightarrow B$ is the bundle projection. From Section 4.1, we know that on this bundle one can choose a connection, which defines a horizontal distribution on TQ (cf. Definition 4.1.3), and associated to such a connection, there is a unique connection form, $\omega \in \varpi^1(Q; \mathcal{G}')$. We now consider a class of constraints which relate to the connection form as follows.

Constraint Hypothesis 5.2.1: We assume that the motion of the system, $(q(\cdot), v_q(\cdot))$, is constrained to a $2n - p'$ dimensional subspace of TQ defined by

$$\mathcal{S} \triangleq \{ (q, v_q) \in TQ \mid \omega(q)(v_q) = \xi(q) \}, \tag{5.2.4}$$

where the mapping $\xi : Q \rightarrow \mathcal{G}'$ is smooth and also G'-equivariant, i.e., $\xi(g \cdot q) = Ad_g \xi(q), \forall g \in G'$. ∎

Remark 5.2.2: Since $\mathcal{G}'$ is a p'-dimensional vector space and $\omega(q)(v_q)$ is linear in v_q, the above assumption can be viewed as giving p' affine constraints on TQ. In addition, if $\xi(q) \equiv 0$, the subspace $\mathcal{S}$ is just the horizontal distribution. ∎

With the above setting, according to equations (2.2.33) derived from Lagrange-d'Alembert principle, the dynamics of the system is given by the following equations,

$$\frac{d}{dt} D_2 L(q, v_q) \cdot u_q - D_1 L(q, v_q) \cdot u_q = F \cdot u_q \tag{5.2.5a}$$

for (q, v_q) satisfying

$$\omega(q)(v_q) = \xi(q) \tag{5.2.5b}$$

and u_q belonging to the horizontal subspace at q, $H_q \subset T_q Q$, i.e.,

$$\omega(q)(u_q) = 0. \qquad (5.2.5c)$$

Here, F is an exterior force satisfying the following hypothesis:

Exterior Force Hypothesis 5.2.3: We assume that an exterior force is a mapping

$$F : TQ \to T^* Q$$

such that $\forall (q, v_q)$,

$$F(\Phi_g(q), T_q \Phi_g \cdot v_q) = T_q^* \Phi_g \cdot F(q, v_q) \qquad \forall g \in G, \qquad (5.2.6)$$

i.e., it is G-equivariant. ∎

Let $q(\cdot) = \{q(t); t \geq 0\}$ be the solution of (5.2.5) with initial condition $q(0) = q_0$, and $x(\cdot) = \{x(t) = \pi(q(t)), t \geq 0\}$ the projection of $q(\cdot)$ on B. Our final goal is to explicitly formulate the projected dynamics on TB. Since we assume that the structure group G' is non-Abelian in general, such a formulation is no longer a straightforward calculation as we have done in the previous section. Our strategy now is to determine the unconstrained dynamics on the horizontal subspace for the given connection, and then, project it to the base space B.

Given a connection on Q, let $r(\cdot) = \{r(t), t \geq 0\}$ be the horizontal lift of $x(\cdot)$ to Q for given $r(0) = r_0, \pi(r_0) = x(0) = \pi(q_0)$. From Theorem 4.1.7, we know that this lifted curve is unique. The question we will address is that if the left-hand-side of Euler-Lagrange equation (5.2.5a) is restricted to the horizontal curve $r(\cdot)$, what the dynamic equation should be, i.e., we will answer the following question:

$$\frac{d}{dt} D_2 L(r, v_r) \cdot u_r - D_1 L(r, v_r) \cdot u_r = ?, \qquad (5.2.7)$$

where $u_r \in H_r$ (the horizontal subspace at r) and $v_r = \dot{r}(t)$.

Remark 5.2.4: Equivalently, one can ask what is the form of equation (5.2.5a) if v_q is split into horizontal and vertical components. ∎

From the uniqueness of horizontal lift, we know that for a given $q(\cdot)$ in Q and the horizontal lift $r(\cdot)$, there exists a unique curve $g(\cdot) = \{g(t), t \geq 0\} \in G'$ such that

$$q(t) = \Phi(g(t), r(t)) \stackrel{\triangle}{=} \Phi_g(r) \stackrel{\triangle}{=} g \cdot r.$$

Then,

$$
\begin{aligned}
v_q(t) = \dot{q}(t) &= T_r\Phi_g \cdot v_r + \left.\frac{d}{d\tau}\right|_{\tau=t} \Phi(g(\tau)g^{-1}(t), g(t) \cdot r(t)) \\
&= T_r\Phi_g \cdot v_r + [T_g R_{g^{-1}}\dot{g}(t)]_Q(q) \\
&= T_r\Phi_g \cdot v_r + [\eta(t)]_Q(q),
\end{aligned}
\tag{5.2.8}
$$

where $\eta(t) = T_g R_{g^{-1}}\dot{g}(t) \in \mathcal{G}'$ for $\mathcal{G}'$ the Lie algebra of G'. Equation (5.2.8) presents the splitting of a tangent vector on Q according to a choice of connection. Evaluate $\omega(q)$ on both side of (5.2.8). Since v_r is horizontal, by Constraint Hypothesis 5.2.1 we have

$$
\eta(t) = \omega(q)(v_q) = \xi(q).
\tag{5.2.9}
$$

Therefore,

$$
v_q(t) = T_r\Phi_g \cdot v_r + [\xi(q)]_Q(q).
\tag{5.2.10}
$$

Since the connection form has equivariant property, we have

$$
\begin{aligned}
v_r &= T_q\Phi_{g^{-1}} \cdot v_q - T_q\Phi_{g^{-1}} \cdot [\xi(q)]_Q(q) \\
&= T_q\Phi_{g^{-1}} \cdot v_q - [Ad_{g^{-1}}\xi(q)]_Q(g^{-1} \cdot q) \\
&= T_q\Phi_{g^{-1}} \cdot v_q - [\xi(r)]_Q(r).
\end{aligned}
\tag{5.2.11}
$$

For simplicity, we sometimes abbreviate $[\xi(\cdot)]_Q(\cdot)$ by $\xi_Q(\cdot)$. But when it is necessary, e.g. operating on derivatives or displaying the final results, we will use the complete notation.

In the following derivations, we will frequently use Equations (5.2.6), (5.2.10), (5.2.11), the G-invariant property of the Riemannian metric and potential energy, and the chain rule in differentiation. The derivations will be carried out in local coordinates, but in an intrinsic way.

Substituting v_r from (5.2.11) into the first term of left-hand-side of (5.2.7), we have

$$
\begin{aligned}
\frac{d}{dt}D_2 L(r, v_r) \cdot u_r &= \frac{d}{dt}K(r)(v_r, u_r) \\
&= \frac{d}{dt}K(r)(T_q\Phi_{g^{-1}} \cdot v_q, u_r) - \frac{d}{dt}K(r)(T_q\Phi_{g^{-1}} \cdot [\xi(q)]_Q(q), u_r) \\
&\triangleq \mathcal{A} - \mathcal{B}.
\end{aligned}
\tag{5.2.12}
$$

Here, $\mathcal{A}$ can be further expanded as

$$
\begin{aligned}
\mathcal{A} &= \frac{d}{dt} K(g \cdot r)(v_q, T_r \Phi_g \cdot u_r) \\
&= \frac{d}{d\tau}\Big|_{\tau=t} K(q(\tau))(v_q(\tau), T_r \Phi_{g(t)} \cdot u_r) + \frac{d}{d\tau}\Big|_{\tau=t} K(q(t))(v_q(t), T_r \Phi_{g(\tau)} \cdot u_r) \\
&\triangleq \mathcal{A}_1 + \mathcal{A}_2.
\end{aligned}
\tag{5.2.13}
$$

Let $u_q = T_r \Phi_g \cdot u_r$ which certainly satisfies condition (5.2.5c). Then, from (5.2.5a), we have

$$
\begin{aligned}
\mathcal{A}_1 &= \frac{1}{2}(D_q K(q) \cdot u_q)(v_q, v_q) - D_q V(q) \cdot u_q + F(q, v_q) \cdot u_q \\
&= \frac{1}{2}(D_q K(g \cdot r) \cdot T_r \Phi_g u_r)(T_r \Phi_g(v_r + \xi_Q(r)), T_r \Phi_g(v_r + \xi_Q(r))) \\
&\quad - (D_q V(g \cdot r) \cdot T_r \Phi_g u_r) + F(g \cdot r, T_q \Phi_g(v_r + [\xi(r)]_Q(r))) \cdot T_r \Phi_g u_r \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} [\frac{1}{2} K(g \cdot (r + \epsilon u_r))(T_r \Phi_g(v_r + \xi_Q(r)), T_r \Phi_g(v_r + \xi_Q(r))) \\
&\quad - V(g \cdot (r + \epsilon u_r))] + T_r^* \Phi_g F(r, v_r + [\xi(r)]_Q(r)) \cdot T_r \Phi_g u_r \\
&= \frac{1}{2}(D_r K(r) \cdot u_r)(v_r, v_r) - D_r V(r) \cdot u_r + F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r \\
&\quad + \frac{1}{2}(D_r K(r) \cdot u_r)(\xi_Q(r), \xi_Q(r)) + (D_r K(r) \cdot u_r)(\xi_Q(r), v_r).
\end{aligned}
\tag{5.2.14}
$$

Note that, in the above derivation, we used the Exterior Force Hypothesis 5.2.3. The first two terms in (5.2.14) is in fact the second term in (5.2.6), that is, $D_1 L(r, v_r) \cdot u_r$.

We now consider $\mathcal{A}_2$ in (5.2.13). Note that

$$
\begin{aligned}
\frac{d}{dt} T_r \Phi_{g(t)} \cdot u_r &= \frac{d}{dt} \frac{d}{d\epsilon}\Big|_{\epsilon=0} \Phi(g(t), r + \epsilon u_r) \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} \frac{d}{d\tau}\Big|_{\tau=t} \Phi(g(\tau)g^{-1}(t), g(t) \cdot (r + \epsilon u_r)) \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} [\xi(g \cdot r)]_Q(g \cdot (r + \epsilon u_r)) \\
&= T_r \Phi_g \cdot \frac{d}{d\epsilon}\Big|_{\epsilon=0} [\xi(r)]_Q(r + \epsilon u_r) \\
&= T_r \Phi_g \cdot (D_r \xi_Q(r) \cdot u_r).
\end{aligned}
\tag{5.2.15}
$$

Therefore,

$$
\begin{aligned}
\mathcal{A}_2 &= K(q)(v_q, \frac{d}{dt} T_r \Phi_{g(t)} \cdot u_r) \\
&= K(g \cdot r)(T_r \Phi_g \cdot (v_r + \xi_Q(r)), T_r \Phi_g \cdot (D_r \xi_Q(r) \cdot u_r)) \\
&= K(r)(v_r, D_r \xi_Q(r) \cdot u_r) + K(r)(\xi_Q(r), D_r \xi_Q(r) \cdot u_r).
\end{aligned}
\tag{5.2.16}
$$

123

Finally, $\mathcal{B}$ in (5.2.12) can be re-arranged as follows:

$$
\begin{aligned}
\mathcal{B} =& \frac{d}{dt} K(q(t))([\xi(q(t))]_Q(q(t)), u_{q(t)}) \\
=& (D_q K(q) \cdot v_q)(\xi_Q(q), u_q) + K(q)([D_q \xi(q) \cdot v_q]_Q(q), u_q) \\
& + K(q)(D_q \xi_Q(q) \cdot v_q, u_q) + K(q)(\xi_Q(q), \frac{d}{dt} T_r \Phi_{g(t)} u_r) \\
\triangleq& \ \mathcal{B}_1 + \mathcal{B}_2 + \mathcal{B}_3 + \mathcal{B}_4,
\end{aligned}
\tag{5.2.17}
$$

where

$$
\begin{aligned}
\mathcal{B}_1 =& (D_q K(g \cdot r) \cdot T_r \Phi_g(v_r + \xi_Q(r)))(T_r \Phi_g \xi_Q(r), T_r \Phi_g u_r) \\
=& \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(g \cdot (r + \epsilon(v_r + \xi_Q(r))))(T_r \Phi_g \xi_Q(r), T_r \Phi_g u_r) \\
=& \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(r + \epsilon(v_r + \xi_Q(r)))(\xi_Q(r), u_r) \\
=& (D_r K(r) \cdot (v_r + \xi_Q(r)))(\xi_Q(r), u_r),
\end{aligned}
\tag{5.2.18}
$$

$$
\begin{aligned}
\mathcal{B}_2 =& K(q)([\frac{d}{d\epsilon}\Big|_{\epsilon=0} \xi(q + \epsilon v_q)]_Q(q), T_r \Phi_g u_r) \\
=& K(g \cdot r)([\frac{d}{d\epsilon}\Big|_{\epsilon=0} \xi(g \cdot (r + \epsilon(v_r + \xi_Q(r))))]_Q(g \cdot r), T_r \Phi_g u_r) \\
=& K(g \cdot r)(T_r \Phi_g[\frac{d}{d\epsilon}\Big|_{\epsilon=0} \xi(r + \epsilon(v_r + \xi_Q(r)))]_Q(r), T_r \Phi_g u_r) \\
=& K(r)([D_r \xi(r) \cdot (v_r + \xi_Q(r))]_Q(r), u_r),
\end{aligned}
\tag{5.2.19}
$$

$$
\begin{aligned}
\mathcal{B}_3 =& K(q)(\frac{d}{d\epsilon}\Big|_{\epsilon=0} [\xi(q)]_Q(q + \epsilon v_q), u_q) \\
=& K(g \cdot r)(\frac{d}{d\epsilon}\Big|_{\epsilon=0} [\xi(g \cdot r)]_Q(g \cdot (r + \epsilon(v_r + \xi_Q(r)))), T_r \Phi_g u_r) \\
=& K(r)(D_r \xi_Q(r) \cdot (v_r + \xi_Q(r)), u_r),
\end{aligned}
\tag{5.2.20}
$$

$$
\begin{aligned}
\mathcal{B}_4 =& K(g \cdot r)(T_r \Phi_g \xi_Q(r), T_r \Phi_g D_r \xi_Q(r) \cdot u_r) \\
=& K(r)(D_r \xi_Q(r) \cdot u_r, \xi_Q(r)).
\end{aligned}
\tag{5.2.21}
$$

From (5.2.12)-(5.2.21), we get the answer to the question in (5.2.7) as follows:

124

$$\frac{d}{dt}D_2 L(r, v_r) \cdot u_r - D_1 L(r, v_r) \cdot u_r = F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r$$
$$+ \frac{1}{2}(D_r K(r) \cdot u_r)(\xi_Q(r), \xi_Q(r))$$
$$+ (D_r K(r) \cdot u_r)(\xi_Q(r), v_r)$$
$$+ K(r)(v_r, D_r \xi_Q(r) \cdot u_r)$$
$$- (D_r K(r) \cdot (v_r + \xi_Q(r)))(\xi_Q(r), u_r)$$
$$- K(r)([D_r \xi(r) \cdot (v_r + \xi_Q(r))]_Q(r), u_r)$$
$$- K(r)(D_r \xi_Q(r) \cdot (v_r + \xi_Q(r)), u_r). \quad (5.2.22)$$

Next, we will make (5.2.22) more compact by defining new functions. Let
$$V^\xi(r) \triangleq V(r) + \frac{1}{2}K(r)([\xi(r)]_Q(r), [\xi(r)]_Q(r))$$
$$= V(r) + \frac{1}{2}\langle \mathbb{I}(r)\xi(r), \xi(r) \rangle, \quad (5.2.23)$$

where the mapping $\mathbb{I}(r) : \mathcal{G} \to \mathcal{G}^*$ has been defined in (4.1.9). Then
$$D_r V^\xi(r) \cdot u_r = D_r V(r) \cdot u_r + \frac{1}{2}(D_r K(r) \cdot u_r)(\xi_Q(r), \xi_Q(r))$$
$$+ K(r)([D_r \xi(r) \cdot u_r]_Q(r), \xi_Q(r)) + K(r)(D_r \xi_Q(r) \cdot u_r, \xi_Q(r)). \quad (5.2.24)$$

Define a new Lagrangian on horizontal space $\mathbf{H}_r$,
$$L^\xi(r, v_r) \triangleq \frac{1}{2}K(r)(v_r, v_r) - V^\xi(r)$$

and a function on TQ:
$$\Xi(r)(u_r, v_r) \triangleq K(r)(v_r, [D_r \xi(r) \cdot u_r]_Q(r)).$$

Then (5.2.22) becomes,
$$\frac{d}{dt}D_2 L^\xi(r, v_r) \cdot u_r - D_1 L^\xi(r, v_r) \cdot u_r = F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r$$
$$+ (D_r K(r) \cdot u_r)(\xi_Q(r), v_r + \xi_Q(r))$$
$$+ K(r)(D_r \xi_Q(r) \cdot u_r, v_r + \xi_Q(r))$$
$$+ K(r)([D_r \xi(r) \cdot u_r]_Q(r), v_r + \xi_Q(r))$$
$$- (D_r K(r) \cdot (v_r + \xi_Q(r)))(\xi_Q(r), u_r)$$
$$- K(r)(D_r \xi_Q(r) \cdot (v_r + \xi_Q(r)), u_r)$$
$$- K(r)([D_r \xi(r) \cdot (v_r + \xi_Q(r))]_Q(r), u_r)$$
$$- \Xi(r)(u_r, v_r)$$

125

$$= F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r + (D_r[K^\flat(r)([\xi(r)]_Q(r))] \cdot u_r) \cdot (v_r + \xi_Q(r))$$

$$- (D_r[K^\flat(r)([\xi(r)]_Q(r))] \cdot (v_r + \xi_Q(r))) \cdot u_r - \Xi(r)(u_r, v_r)$$

$$= F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r + d\omega_\xi(r)(u_r, v_r + \xi_Q(r)) - \Xi(r)(u_r, v_r),$$

where $\omega_\xi(r) \triangleq K^\flat(r)([\xi(r)]_Q(r))$ is a real-valued one-form on Q and $K^\flat : TQ \to T^*Q$ is the usual Legendre transform. In the above derivation, we used the following fact in local coordinates,

$$d\alpha(r)(X, Y) = (D\alpha(r) \cdot X) \cdot Y - (D\alpha(r) \cdot Y) \cdot X \quad \forall r \in Q$$

for $X, Y \in \mathfrak{X}(Q)$ and any one-form α on Q.

Up to now, we have proved the following theorem.

Theorem 5.2.5: If $q(\cdot)$ is a solution of the constrained dynamics (5.2.5), then, for given choice of connection on principal G'-bundle (5.2.3), any horizontal lift, $r(\cdot)$, of $q(\cdot)$'s projection satisfies the unconstrained equation

$$\frac{d}{dt} D_2 L^\xi(r, v_r) \cdot u_r - D_1 L^\xi(r, v_r) \cdot u_r = F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r$$
$$+ d\omega_\xi(r)(u_r, v_r) + d\omega_\xi(r)(u_r, [\xi(r)]_Q(r))$$
$$- \Xi(r)(u_r, v_r) \tag{5.2.25}$$

for any $u_r \in \mathbf{H}_r \subset T_r Q$, where

$$L^\xi(r, v_r) = \frac{1}{2} K(r)(v_r, v_r) - V^\xi(r)$$

$$V^\xi(r) = V(r) + \frac{1}{2} K(r)([\xi(r)]_Q(r), [\xi(r)]_Q(r))$$

$$= V(r) + \frac{1}{2} \langle \mathbb{I}(r)\xi(r), \xi(r) \rangle,$$

$$\omega_\xi(r) = K^\flat(r)([\xi(r)]_Q(r)),$$

$$\Xi(r)(u_r, v_r) = K(r)(v_r, [D_r\xi(r) \cdot u_r]_Q(r)). \qquad \blacksquare$$

Remarks 5.2.6:

(1) From the derivation of $d\omega_\xi$, the force

$$F_{gyro}(r, v_r) \triangleq d\omega_\xi(r)(\,\cdot\,, v_r) \tag{5.2.26}$$

has the property of a gyroscopic force.

(2) One can show that Equation (5.2.22) can be written in an alternative form without resort to the function V^ξ, that is,

$$
\begin{aligned}
\frac{d}{dt} D_2 L(r, v_r) \cdot u_r - D_1 L(r, v_r) \cdot u_r ={} & F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r \\
& + \frac{1}{2} D_r \langle \mathbb{I}(r)\xi(r), \xi(r) \rangle \cdot u_r \\
& + d\omega_\xi(r)(u_r, v_r) - \Xi(r)(u_r, v_r + \xi_Q(r)).
\end{aligned}
$$

We are not going to use this form since, when we consider Lagrangian reduction later, Equation (5.2.25) will give us a straightforward answer.

(3) Once a horizontal curve is determined by solving the unconstrained equation (5.2.25) for an initial condition $r(0)$, the solution for the original constrained equations (5.2.5) can be determined by first solving the differential equation

$$
\dot{g}(t) = g(t) \cdot \xi(r(t))
$$

for $g(0)$ satisfying $q(0) = g(0) \cdot r(0)$, and then setting

$$
q(t) = g(t) \cdot r(t).
$$

(4) The equilibria of the system are determined by the solution of algebraic equations

$$
D_1 V^\xi(r) - d\omega_\xi(r)(\; \cdot \; , [\xi(r)]_Q(r)) = 0. \qquad \blacksquare
$$

5.3   Non-Abelian Chaplygin Systems

In this section, we show how to drop the unconstrained dynamics on horizontal bundle given in (5.2.25) down to the base space for a given principal fiber bundle. To get explicit expressions, we will consider the formulation on product bundles. Since a principal fiber bundle is locally trivial, the following results will be true locally in general. Also, for simplicity, we will assume the symmetry group in (5.2.1) is the same as the structure group of the principal bundle (5.2.3) (i.e., $G' = G$).

Let $Q = B \times G$ be the configuration space parametrized by $q = (x, h)$ for $x \in B$ and $h \in G$. Then, the tangent space is $T_q Q = T_x B \times T_h G$. The tangent vector at any point q in Q is given by

$$
v_q = [\; v_x \; , \; h \cdot \zeta \;]_{(x,h)} \qquad (5.3.1)
$$

for some $\zeta \in \mathcal{G}$, where $h \cdot \zeta \triangleq T_e L_h \zeta$. Consider the principal bundle $(B \times G, B, G)$, where G acts on $Q = B \times G$ on the left as shown in Subsection 4.1.2, and a connection given by connection form $\omega \in \varpi^1(Q; \mathcal{G})$. One can show that, on this bundle there exists an unique $\mathcal{G}$-valued one-form, $\widetilde{\omega}$, on B such that, at each point $q = (x, h)$ in Q,

$$\omega(q) \cdot v_q = Ad_h(\zeta + \widetilde{\omega}(x) \cdot v_x), \tag{5.3.2}$$

where v_q is given in (5.3.1) [16]. We refer to $\widetilde{\omega}$ as the *pull-down connection form*, which is also the usual notion of local connection form if a non-trivial principal bundle is considered. The proof of (5.3.2) can be verified by letting v_q be horizontal and comparing (5.3.2) with (4.1.6). With the above connection, the tangent vector v_q on Q at $q = (x, h)$ has its horizontal and vertical splitting,

$$v_q = Ver(v_q) + Hor(v_q),$$

where

$$Ver(v_q) = [Ad_h(\zeta + \widetilde{\omega}(x) \cdot v_x))]_Q(q) = [\ 0\ ,\ h \cdot (\zeta + \widetilde{\omega}(x) \cdot v_x)\ ] \tag{5.3.3}$$

and

$$Hor(v_q) = v_q - Ver(v_q) = [v_x, -h \cdot (\widetilde{\omega}(x) \cdot v_x)]. \tag{5.3.4}$$

Indeed, (5.3.3) can be carried out directly as follows,

$$\begin{aligned}
Ver(v_q) &= [Ad_h(\zeta + \widetilde{\omega}(x) \cdot v_x)]_Q(q) \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} \Phi(exp\{\epsilon Ad_h(\zeta + \widetilde{\omega}(x) \cdot v_x)\}, (x, h)) \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} [x, exp\{\epsilon Ad_h(\zeta + \widetilde{\omega}(x) \cdot v_x)\} \cdot h] \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} [x, h \cdot exp\{\epsilon(\zeta + \widetilde{\omega}(x) \cdot v_x)\}] \\
&= [0, h \cdot (\zeta + \widetilde{\omega}(x) \cdot v_x)].
\end{aligned}$$

Now we express each object in Theorem 5.2.5 based on the above setting. First, we introduce the induced Riemannian metric and potential function on base space B in accordance with the restriction of Riemannian metric and potential function of a simple mechanical system with symmetry on the horizontal subspace. Let $\mathbf{H}_r$ be the horizontal subspace of $T_r Q$ at $r = (x, h)$ with respect to the above connection. If

$$v_r = [v_x, -h \cdot (\widetilde{\omega}(x) \cdot v_x)] \in \mathbf{H}_r \quad \text{and} \quad w_r = [w_x, -h \cdot (\widetilde{\omega}(x) \cdot w_x)] \in \mathbf{H}_r,$$

then

$$K(r)(v_r, w_r) = K(x, h)([v_x, -h \cdot (\widetilde{\omega}(x) \cdot v_x)], [w_x, -h \cdot (\widetilde{\omega}(x) \cdot w_x)])$$

$$= K(x, e)([v_x, -\widetilde{\omega}(x) \cdot v_x], [w_x, -\widetilde{\omega}(x) \cdot w_x])$$

$$\stackrel{\triangle}{=} \widetilde{K}(x)(v_x, w_x), \tag{5.3.5}$$

where e is the identity element in G, and the G-invariance of K has been used. Also, since the potential function on Q is G-invariant,

$$V(r) = V(x, h) = V(x, e) \stackrel{\triangle}{=} \widetilde{V}(x). \tag{5.3.6}$$

The constraint in Constraint Hypothesis 5.2.1 based on the connection given in (5.3.2) now has the following expression

$$\omega(x, h) \cdot v_{(x,h)} = Ad_h(\zeta + \widetilde{\omega}(x) \cdot v_x) = \xi(x, h) = Ad_h \xi(x, e)$$

or equivalently,

$$\widetilde{\xi}(x) \stackrel{\triangle}{=} \xi(x, e) = \zeta + \widetilde{\omega}(x) \cdot v_x. \tag{5.3.7}$$

Then, for any $q = (x, h) \in Q$,

$$[\xi(q)]_Q(q) = [Ad_h \widetilde{\xi}(x)]_Q(q) = [0, h \cdot \widetilde{\xi}(x)], \tag{5.3.8}$$

and then, the second term of V^ξ in (5.2.23) at q becomes

$$\frac{1}{2} K(q)([\xi(q)]_Q(q), [\xi(q)]_Q(q)) = \frac{1}{2} K(x, h)([0, h \cdot \widetilde{\xi}(x)], [0, h \cdot \widetilde{\xi}(x)])$$

$$= \frac{1}{2} K(x, e)([0, \widetilde{\xi}(x)], [0, \widetilde{\xi}(x)])$$

$$= \frac{1}{2} \langle \widetilde{\mathbb{I}}(x) \widetilde{\xi}(x), \widetilde{\xi}(x) \rangle,$$

where $\widetilde{\mathbb{I}}(x) \stackrel{\triangle}{=} \mathbb{I}(x, e)$.

Using the above newly defined objects on base space B, the function L^ξ in Theorem 5.2.4 can be expressed as,

$$L^\xi(r, v_r) = \frac{1}{2} \widetilde{K}(x)(v_x, v_x) - (\widetilde{V}(x) + \frac{1}{2} \langle \widetilde{\mathbb{I}}(x) \widetilde{\xi}(x), \widetilde{\xi}(x) \rangle)$$

$$\stackrel{\triangle}{=} \widetilde{L}^\xi(x, v_x), \tag{5.3.9}$$

for any point $r = (x, h) \in Q$ and $v_r = [v_x, -h \cdot (\widetilde{\omega}(x) \cdot v_x)] \in \mathrm{H}_r \subset T_r Q$.

129

We are now ready to drop the dynamics given in (5.2.25) down to the base space B. Let $q(\cdot) = (x(\cdot), h(\cdot))$ be the solution of (5.2.5) with initial condition $q(0) = (x_0, h_0)$ for $x_0 \in B$ and $h_0 \in G$, where $x(\cdot)$ is naturally the projection of $q(\cdot)$. Let $r(\cdot) = (x(\cdot), g(\cdot))$ be the horizontal lift of $x(\cdot)$ with initial condition $r(0) = (x(0), g(0))$, where $g(\cdot)$ is determined by solving differential equation

$$\dot{g}(t) = -g(t) \cdot (\widetilde{\omega}(x(t)) \cdot v_x(t)) \tag{5.3.10}$$

for initial condition $g(0)$. Then from (5.3.4),

$$v_r(t) = [v_x(t), -g(t) \cdot (\widetilde{\omega}(x(t)) \cdot v_x(t))]. \tag{5.3.11}$$

Remark 5.3.1: From connection theory, we know that, given a connection on a principal fiber bundle $\wp = (Q, B, \pi, G)$, there exists an unique horizontal lift mapping at each point q in Q, denoted by

$$\mathcal{H}(q) : T_{\pi(q)} \to T_q Q.$$

For trivial bundle case, this mapping has been given explicitly in (5.3.4), i.e.,

$$v_r = \mathcal{H}(x, g) \cdot v_x = [v_x(t), -g(t) \cdot (\widetilde{\omega}(x(t)) \cdot v_x(t))].$$

We will use this notation in the next section. ∎

Let u_r be any tangent vector at $r(t)$ in $\mathbf{H}_r$, which can be represented by

$$u_r = [u_x, -g(t) \cdot (\widetilde{\omega}(x(t)) \cdot u_x)]$$

for some $u_x \in T_x B$. Now, the first term on the left-hand-side of (5.2.25) is

$$\frac{d}{dt} D_2 L^\xi(r, v_r) u_r = \frac{d}{dt} K(r(t))(v_r(t), u_r)$$

$$= \frac{d}{d\tau}\bigg|_{\tau = t} K(x(\tau), g(\tau))([v_x(\tau), -g(\tau) \cdot (\widetilde{\omega}(x(\tau)) v_x(\tau))], [u_x, -g(t) \cdot (\widetilde{\omega}(x(t)) u_x)])$$

$$= \frac{d}{dt} K(x(t), g(t))([v_x(t), -g(t) \cdot (\widetilde{\omega}(x(t)) v_x(t))], [u_x, -g(t) \cdot (\widetilde{\omega}(x(t)) u_x)])$$

$$\quad - K(x, g)([v_x, -g \cdot (\widetilde{\omega}(x) v_x)], \frac{d}{dt}[u_x, -g(t) \cdot (\widetilde{\omega}(x(t)) u_x)])$$

$$= \frac{d}{dt} K(x(t), e)([v_x(t), -\widetilde{\omega}(x(t)) v_x(t)], [u_x, -\widetilde{\omega}(x(t)) u_x])$$

$$\quad - K(x, g)([v_x, -g \cdot (\widetilde{\omega}(x) v_x)], [0, g \cdot (\widetilde{\omega}(x) v_x)(\widetilde{\omega}(x) u_x) - g \cdot ((D_x\widetilde{\omega}(x) v_x) u_x)])$$

$$= \frac{d}{dt} \widetilde{K}(x)(v_x(t), u_x)$$

$$\quad - K(x, e)([v_x, -\widetilde{\omega}(x) v_x], [0, (\widetilde{\omega}(x) v_x)(\widetilde{\omega}(x) u_x) - (D_x\widetilde{\omega}(x) v_x) u_x]). \tag{5.3.12}$$

130

Remark 5.3.2: In the above derivation we used the following convention,

$$g\eta_1\eta_2 \triangleq \frac{d}{d\epsilon}\Big|_{\epsilon=0}\frac{d}{d\lambda}\Big|_{\lambda=0} g \cdot e^{\epsilon\eta_1} \cdot e^{\lambda\eta_2}, \qquad \forall \eta_1, \eta_2 \in \mathcal{G}, \quad \forall g \in G. \qquad \blacksquare$$

The second term in the left-hand-side of (5.2.25) is

$$D_1 L^\xi(r, v_r) \cdot u_r = \frac{1}{2}\left(D_r K(r) \cdot u_r\right)(v_r, v_r) - D_r V(r) \cdot u_r$$
$$- \frac{1}{2} D_r(K(r)([\xi(r)]_Q(r), [\xi(r)]_Q(r))) \cdot u_r, \qquad (5.3.13)$$

where

$$(D_r K(r) \cdot u_r)(v_r, v_r) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(r + \epsilon u_r)(v_r, v_r)$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon u_x, ge^{-\epsilon\widetilde{\omega}(x)u_x})([v_x, -g(\widetilde{\omega}(x)v_x)], [v_x, -g(\widetilde{\omega}(x)v_x)])$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon u_x, e)([v_x, -e^{\epsilon\widetilde{\omega}(x)u_x}(\widetilde{\omega}(x)v_x)], [v_x, -e^{\epsilon\widetilde{\omega}(x)u_x}(\widetilde{\omega}(x)v_x)])$$
$$= (D_x\widetilde{K}(x)u_x)(v_x, v_x) - 2K(x, e)([v_x, -\widetilde{\omega}(x)v_x], [0, -(D_x\widetilde{\omega}(x)u_x)v_x])$$
$$+ 2K(x, e)([v_x, -\widetilde{\omega}(x)v_x], [0, -(\widetilde{\omega}(x)u_x)(\widetilde{\omega}(x)v_x)]), \qquad (5.3.14)$$

$$D_r V(r) \cdot u_r = \frac{d}{d\epsilon}\Big|_{\epsilon=0} V(r + \epsilon u_r)$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} V(x + \epsilon u_x, ge^{-\epsilon\widetilde{\omega}(x)u_x})$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} V(x + \epsilon u_x, e)$$
$$= D_x\widetilde{V}(x) \cdot u_x \qquad (5.3.15)$$

and

$$D_r(K(r)([\xi(r)]_Q(r), [\xi(r)]_Q(r))) \cdot u_r$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(r_\epsilon)([\xi(r_\epsilon)]_Q(r_\epsilon), [\xi(r_\epsilon)]_Q(r_\epsilon))|_{r_\epsilon = r + \epsilon u_r}$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon u_x, ge^{-\epsilon\widetilde{\omega}(x)u_x})([0, ge^{-\epsilon\widetilde{\omega}(x)u_x}\widetilde{\xi}(x + \epsilon u_x)], [0, ge^{-\epsilon\widetilde{\omega}(x)u_x}\widetilde{\xi}(x + \epsilon u_x)])$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon u_x, e)([0, \widetilde{\xi}(x + \epsilon u_x)], [0, \widetilde{\xi}(x + \epsilon u_x)])$$
$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} \langle \widetilde{\mathbb{I}}(x + \epsilon u_x)\widetilde{\xi}(x + \epsilon u_x), \widetilde{\xi}(x + \epsilon u_x)\rangle$$
$$= \langle(D_x\widetilde{\mathbb{I}}(x) \cdot u_x)\widetilde{\xi}(x), \widetilde{\xi}(x)\rangle + 2\langle\widetilde{\mathbb{I}}(x)\widetilde{\xi}(x), (D_x\widetilde{\xi}(x)) \cdot u_x\rangle. \qquad (5.3.16)$$

131

Therefore, from (5.3.12)-(5.3.16), the left-hand-side of (5.2.25) is

$$\frac{d}{dt} D_2 L^\xi(r, v_r) \cdot u_r - D_1 L^\xi(r, v_r) \cdot u_r$$

$$= \frac{d}{dt} \widetilde{K}(x(t))(v_x(t), u_x) - \frac{1}{2}(D_x \widetilde{K}(x)u_x)(v_x, v_x) + D_x \widetilde{V}(x) \cdot u_x$$

$$+ \frac{1}{2} \langle (D_x \widetilde{\mathbb{I}}(x) \cdot u(x))\widetilde{\xi}(x), \widetilde{\xi}(x)\rangle + \langle \widetilde{\mathbb{I}}(x)\widetilde{\xi}(x), (D_x \widetilde{\xi}(x)) \cdot u_x\rangle$$

$$- K(x, e)([v_x, -\widetilde{\omega}(x)v_x], [0, (D_x\widetilde{\omega}(x)u_x)v_x - (D_x\widetilde{\omega}(x)v_x)u_x$$

$$+ (\widetilde{\omega}(x)v_x)(\widetilde{\omega}(x)u_x) - (\widetilde{\omega}(x)u_x)(\widetilde{\omega}(x)v_x)])$$

$$= \frac{d}{dt} D_2 \widetilde{L}^\xi(x, v_x) \cdot u_x - D_1 \widetilde{L}^\xi(x, v_x) \cdot u_x$$

$$- K(x, e)([v_x, -\widetilde{\omega}(x)v_x], [0, \widetilde{\Omega}(x)(u_x, v_x)])$$

$$= \frac{d}{dt} D_2 \widetilde{L}^\xi(x, v_x) \cdot u_x - D_1 \widetilde{L}^\xi(x, v_x) \cdot u_x - \Gamma(x)(v_x, u_x, v_x), \tag{5.3.17}$$

where

$$\widetilde{\Omega}(x)(u_x, v_x) \stackrel{\triangle}{=} (D_x\widetilde{\omega}(x)u_x) \cdot v_x - (D_x\widetilde{\omega}(x)v_x) \cdot u_x$$

$$+ (\widetilde{\omega}(x) \cdot v_x)(\widetilde{\omega}(x) \cdot u_x) - (\widetilde{\omega}(x) \cdot u_x)(\widetilde{\omega}(x) \cdot v_x)$$

$$= d\widetilde{\omega}(x)(u_x, v_x) - [\widetilde{\omega}(x) \cdot u_x, \widetilde{\omega}(x) \cdot v_x] \tag{5.3.18}$$

and

$$\Gamma(x)(v_x, u_x, v_x) \stackrel{\triangle}{=} K(x, e)([v_x, -\widetilde{\omega}(x) \cdot v_x], [0, \widetilde{\Omega}(x)(u_x, v_x)]). \tag{5.3.19}$$

Among the newly defined objects above, $\Gamma(x)(v_x, u_x, v_x)$ is a $(3, 0)$-tensor on B, which is skew-symmetric in last two vectors. And, $\widetilde{\Omega}$ is a $\mathcal{G}$-valued two-form on B, which plays the role of local curvature form. Indeed, we have following result.

Proposition 5.3.3: Let $\bar{r} = (\bar{x}, \bar{g})$ be any point in Q for any point $\bar{x} \in B$ and $\bar{g} \in G$, and

$$u_{\bar{r}} = (u_{\bar{x}}, -\bar{g} \cdot (\widetilde{\omega}(\bar{x}) \cdot u_{\bar{x}})) \quad \text{and} \quad v_{\bar{r}} = (v_{\bar{x}}, -\bar{g} \cdot (\widetilde{\omega}(\bar{x}) \cdot v_{\bar{x}}))$$

be horizontal tangent vectors with respect to the connection given in (5.3.2). Let Ω be the corresponding curvature form of the connection. Then we have

$$[\Omega(\bar{r})(u_{\bar{r}}, v_{\bar{r}})]_Q(\bar{r}) = (0, \bar{g} \cdot \widetilde{\Omega}(\bar{x})(u_{\bar{x}}, v_{\bar{x}})). \tag{5.3.20}$$

Proof: Recall that, if U and V are horizontal vector fields on Q,

$$[\Omega(q)(U(q), V(q))]_Q(q) = -Ver([U, V](q)). \tag{5.3.21}$$

132

We start from computing $[\,U, V\,]$. Let $u(x)$ and $v(x)$ be vector fields on B extended respectively from $u_{\bar{r}}$ and $v_{\bar{x}}$, i.e., $u(\bar{x}) = u_{\bar{x}}$ and $v(\bar{x}) = v_{\bar{x}}$. Then we define two horizontal vector fields on Q to be, for any $r = (x, g) \in Q$,

$$U(r) = [u(x), -g \cdot (\widetilde{\omega}(x) \cdot u(x))] \quad \text{and} \quad V(r) = [v(x), -g \cdot (\widetilde{\omega}(x) \cdot v(x))]. \quad (5.3.22)$$

Let $\phi_U(t, r)$ and $\phi_V(t, r)$ be the tangent curves of U and V at r, respectively, i.e.,

$$\phi_U(t, r) = (x + tu(x), g \cdot e^{-t\widetilde{\omega}(x)u(x)}) \quad \text{and} \quad \phi_V(t, r) = (x + tv(x), g \cdot e^{-t\widetilde{\omega}(x)v(x)})$$

Then, using Lemma 4.4.1,

$$
\begin{aligned}
[\,U, V\,](r) &= \frac{d}{dt}\Big|_{t=0} V(\phi_U(t, r)) - U(\phi_V(t, r)) \\
&= \frac{d}{dt}\Big|_{t=0} [v(x + tu(x)), -ge^{-t\widetilde{\omega}(x)u(x)}\widetilde{\omega}(x + tu(x))v(x + tu(x))] \\
&\quad - \frac{d}{dt}\Big|_{t=0} [u(x + tv(x)), -ge^{-t\widetilde{\omega}(x)v(x)}\widetilde{\omega}(x + tv(x))u(x + tv(x))] \\
&= [Dv(x)u(x), g(\widetilde{\omega}(x)u(x))(\widetilde{\omega}(x)v(x)) - g(D\widetilde{\omega}(x)u(x))v(x) - g\widetilde{\omega}(x)(Dv(x)u(x))] \\
&\quad - [Du(x)v(x), g(\widetilde{\omega}(x)v(x))(\widetilde{\omega}(x)u(x)) + g(D\widetilde{\omega}(x)v(x))u(x) + g\widetilde{\omega}(x)(Du(x)v(x))] \\
&= [Dv(x)u(x) - Du(x)v(x), -g\widetilde{\omega}(x)(Dv(x)u(x) - Du(x)v(x))] \\
&\quad + [0, g((\widetilde{\omega}(x)u(x))(\widetilde{\omega}(x)v(x)) - (\widetilde{\omega}(x)v(x))(\widetilde{\omega}(x)u(x)) \\
&\qquad\qquad + (D\widetilde{\omega}(x)v(x))u(x) - (D\widetilde{\omega}(x)u(x))v(x))] \\
&= [Dv(x)u(x) - Du(x)v(x), -g \cdot \widetilde{\omega}(x)(Dv(x)u(x) - Du(x)v(x))] \\
&\quad + [0, g \cdot ([\,\widetilde{\omega}(x)u(x), \widetilde{\omega}(x)v(x)\,] - d\widetilde{\omega}(x)(u(x), v(x)))]. \quad (5.3.23)
\end{aligned}
$$

It is clear that in (5.3.23) the first term is the horizontal part of $[\,U, V\,](r)$ and the second term is vertical part of $[\,U, V\,](r)$ which, at $(\bar{x}, \bar{g})$, is identical to the minus of right-hand-side of (5.3.20). ∎

Remark 5.3.4: If one assumes G acts on Q to the *right* in the definition of principal fiber bundle, $\widetilde{\Omega}$ is of the form

$$\widetilde{\Omega}(x)(u_x, v_x) = d\widetilde{\omega}(x)(u_x, v_x) + [\,\widetilde{\omega}(x)(u_x), \widetilde{\omega}(x)(v_x)\,],$$

which is called local curvature form in many textbooks in physics and mathematics. ∎

We now start to drop the right-hand-side of (5.2.25) down to the base space B. The first term there can written as, using Exterior Force Hypothesis 5.2.3,

$$
\begin{aligned}
F(r, v_r + [\xi(r)]_Q(r)) \cdot u_r &= F((x, g), [v_x, g(\widetilde{\xi}(x) - \widetilde{\omega}(x) \cdot v_x]) \cdot [u_x, -g\widetilde{\omega}(x) \cdot u_x] \\
&= F((x, e), [v_x, \widetilde{\xi}(x) - \widetilde{\omega}(x) \cdot v_x]) \cdot [u_x, -\widetilde{\omega}(x) \cdot u_x] \\
&\stackrel{\triangle}{=} \widetilde{F}(x, v_x) \cdot u_x.
\end{aligned}
\tag{5.3.24}
$$

We give the drop of the second and third terms on the right-hand-side of (5.2.25) in the following Lemmas.

Lemma 5.3.5: Let $\bar{r} = (\bar{x}, \bar{g})$ be any point in Q for any point $\bar{x} \in B$ and $\bar{g} \in G$, and

$$
u_{\bar{r}} = (u_{\bar{x}}, -\bar{g} \cdot (\widetilde{\omega}(\bar{x}) \cdot u_{\bar{x}})) \quad \text{and} \quad v_{\bar{r}} = (v_{\bar{x}}, -\bar{g} \cdot (\widetilde{\omega}(\bar{x}) \cdot v_{\bar{x}}))
$$

be horizontal tangent vectors at $\bar{r}$ with respect to the connection given in (5.3.2). Then,

$$
d\omega_\xi(\bar{r})(u_{\bar{r}}, v_{\bar{r}}) = d\widetilde{\omega}_\xi(\bar{x})(u_{\bar{x}}, v_{\bar{x}}) + \langle \widetilde{\mathbb{I}}(\bar{x})\widetilde{\xi}(\bar{x}), \widetilde{\Omega}(\bar{x})(u_{\bar{x}}, v_{\bar{x}}) \rangle,
\tag{5.3.25}
$$

where $\widetilde{\omega}_\xi$ is a one-form on B defined by, for any $z \in B$ and $w_z \in T_z B$,

$$
\widetilde{\omega}_\xi(z)(w_z) \stackrel{\triangle}{=} K(z, e)([0, \widetilde{\xi}(z)], [w_z, -\widetilde{\omega}(z)w_z]).
\tag{5.3.26}
$$

Proof:

As we have done in the proof of Proposition 5.3.3, let U and V be the vector fields extended from tangent vectors $u_{\bar{r}}$ and $v_{\bar{r}}$, respectively, such that at any point r in Q, their expressions are given in (5.3.22). From (4.1.3),

$$
d\omega_\xi(r)(U(r), V(r)) = U(r)[\omega_\xi(r)V(r)] - V(r)[\omega_\xi(r)U(r)] - \omega_\xi(r)([U, V](r)).
\tag{5.3.27}
$$

From the definition of ω_ξ in Theorem 5.2.5,

$$
\begin{aligned}
\omega_\xi(r)V(r) &= K(r)([\xi(r)]_Q(r), V(r)) \\
&= K(x, g)([0, g \cdot \widetilde{\xi}(x)], [v(x), -g \cdot (\widetilde{\omega}(x) \cdot v(x))]).
\end{aligned}
$$

Then, letting $\phi_U(\epsilon, r) = [x + \epsilon u(x), ge^{-\epsilon\widetilde{\omega}(x)u(x)}]$ be the tangent curve of U at $r = (x, g)$, we have

$$
U(r)[\omega_\xi(r)V(r)] = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \omega_\xi(\phi_U(\epsilon, r))V(\phi_U(\epsilon, r))
$$

134

$$= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} K(x + \epsilon u(x), ge^{-\epsilon\widetilde{\omega}(x)u(x)})([0, ge^{-\epsilon\widetilde{\omega}(x)u(x)}\widetilde{\xi}(x + \epsilon u(x))],$$

$$[v(x + \epsilon u(x)), -ge^{-\epsilon\widetilde{\omega}(x)u(x)}\widetilde{\omega}(x + \epsilon u(x))v(x + \epsilon u(x))])$$

$$= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} K(x + \epsilon u(x), e)([0, \widetilde{\xi}(x + \epsilon u(x))],$$

$$[v(x + \epsilon u(x)), -\widetilde{\omega}(x + \epsilon u(x))v(x + \epsilon u(x))])$$

$$= D(\widetilde{\omega}_\xi(x) \cdot u(x)) \cdot v(x)$$

$$+ K(x, e)([0, \widetilde{\xi}(x)], [Dv(x) \cdot u(x), -\widetilde{\omega}(x)(Dv(x) \cdot u(x))]). \qquad (5.3.28)$$

Following the same way, we get

$$V(r)[\omega_\xi(r)U(r)] = D(\widetilde{\omega}_\xi(x) \cdot v(x)) \cdot u(x)$$

$$+ K(x, e)([0, \widetilde{\xi}(x)], [Du(x) \cdot v(x), -\widetilde{\omega}(x)(Du(x) \cdot v(x))]). \qquad (5.3.29)$$

The Lie bracket of vector field U and V has been computed in (5.3.23). Then

$$\omega_\xi(r)([\,U, V\,](r))$$

$$= K(x, e)([0, \widetilde{\xi}(x)],$$

$$[Dv(x) \cdot u(x) - Du(x) \cdot v(x), -\widetilde{\omega}(x)(Dv(x) \cdot u(x) - Du(x) \cdot v(x))])$$

$$+ K(x, e)([0, \widetilde{\xi}(x)], [0, \widetilde{\Omega}(x)(u(x), v(x))]). \qquad (5.3.30)$$

Substituting (5.3.28)-(5.3.30) into (5.3.27), we get

$$d\omega_\xi(r)(U(r), V(r)) = d\widetilde{\omega}_\xi(x)(u(x), v(x)) + \langle\widetilde{\mathbb{I}}(x)\widetilde{\xi}(x), \widetilde{\Omega}(x)(u(x), v(x))\rangle.$$

Evaluating the above equation at $r = \bar{r}$, we get (5.3.25). ∎

Lemma 5.3.6: Let $\bar{r}$ and $u_{\bar{r}}$ be defined as in Lemma 5.3.5. Let $[\xi(\bar{r})]_Q(\bar{r}) = [0, \bar{g}\cdot\widetilde{\xi}(\bar{x})]$. Then,

$$d\omega_\xi(\bar{r})(u_{\bar{r}}, [\xi(\bar{r})]_Q(\bar{r})) = \langle(D_{\bar{x}}\widetilde{\mathbb{I}}(\bar{x}) \cdot u_{\bar{x}})\widetilde{\xi}(\bar{x}), \widetilde{\xi}(\bar{x})\rangle + \langle\widetilde{\mathbb{I}}(\bar{x})\widetilde{\xi}(\bar{x}), D_x\widetilde{\xi}(\bar{x}) \cdot u_{\bar{x}}\rangle$$

$$+ \langle\,\widetilde{\mathbb{I}}(\bar{x})\widetilde{\xi}(\bar{x}), [\widetilde{\omega}(\bar{x})u_{\bar{x}}, \widetilde{\xi}(\bar{x})]\,\rangle. \qquad (5.3.31)$$

Proof:

Let U be the same vector field given in (5.3.22). Let $\bar{\xi}_Q$ be the vector field extended from $\bar{\xi}_Q(\bar{x})$ such that, at any point r in Q,

$$\bar{\xi}_Q(r) = [Ad_{\bar{g}}\widetilde{\xi}(\bar{x})]_Q(x, g) = [0, Ad_{\bar{g}}\widetilde{\xi}(\bar{x}) \cdot g].$$

Again, from (4.1.3), we have

$$d\omega_\xi(r)(U(r),\bar\xi_Q(r)) = U(r)[\omega_\xi(r)\bar\xi_Q(r)] - \bar\xi_Q(r)[\omega_\xi(r)U(r)] - \omega_\xi(r)([U,\bar\xi_Q](r)). \quad (5.3.32)$$

Let

$$\phi_U(\epsilon,r) = [x + \epsilon u(x), ge^{-\epsilon\widetilde\omega(x)u(x)}] \quad \text{and} \quad \phi_{\bar\xi_Q}(\epsilon,r) = [x, e^{\epsilon Ad_{\bar g}\widetilde\xi(\bar x)}g]$$

be tangent curves of vector fields of U and $\bar\xi_Q$ at (x,g), respectively. Then following the same procedure as we have done in the proof of previous Lemma, we have

$$U(r)[\omega_\xi(r)(\bar\xi_Q(r))] = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \omega_\xi(\phi_U(\epsilon,r))(\bar\xi_Q(\phi_U)(\epsilon,r))$$

$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon u(x), ge^{-\epsilon\widetilde\omega(x)u(x)})([0, ge^{-\epsilon\widetilde\omega(x)u(x)}\widetilde\xi(x + \epsilon u(x))],$$

$$[0, Ad_{\bar g}\widetilde\xi(\bar x)ge^{-\epsilon\widetilde\omega(x)u(x)}])$$

$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon u(x), e)([0, \widetilde\xi(x + \epsilon u(x))],$$

$$[0, e^{\epsilon\widetilde\omega(x)u(x)}g^{-1}Ad_{\bar g}\widetilde\xi(\bar x)ge^{-\epsilon\widetilde\omega(x)u(x)}])$$

$$= \langle D\widetilde{\mathbb{I}}(x)\widetilde\xi(x), Ad_{g^{-1}\bar g}\widetilde\xi(\bar x)\rangle + \langle \widetilde{\mathbb{I}}(x)D\widetilde\xi(x)\cdot u(x), Ad_{g^{-1}\bar g}\widetilde\xi(\bar x)\rangle$$

$$+ \langle \widetilde{\mathbb{I}}(x)\widetilde\xi(x), (\widetilde\omega(x)u(x))Ad_{g^{-1}\bar g}\widetilde\xi(\bar x) - Ad_{g^{-1}\bar g}\widetilde\xi(\bar x)(\widetilde\omega(x)u(x))\rangle \quad (5.3.33)$$

and

$$\bar\xi_Q(r)[\omega_\xi(r)U(r)] = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \omega_\xi(\phi_{\bar\xi_Q}(\epsilon,r))U(\phi_{\bar\xi_Q}(\epsilon,r))$$

$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x, e^{\epsilon Ad_{g^{-1}}\widetilde\xi(\bar x)}g)([0, e^{\epsilon Ad_{g^{-1}}\widetilde\xi(\bar x)}g\widetilde\xi(x)], [u(x), -e^{\epsilon Ad_{g^{-1}}\widetilde\xi(\bar x)}g\widetilde\omega(x)u(x)])$$

$$= 0. \quad (5.3.34)$$

Finally, using Lemma 4.4.2, we have

$$[U,\bar\xi_Q](r) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \bar\xi_Q(\phi_U(\epsilon,r)) - U(\phi_{\bar\xi_Q}(\epsilon,r)).$$

But

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} \bar\xi_Q(\phi_U(\epsilon)) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} [0, Ad_{g^{-1}}\widetilde\xi(\bar x)ge^{-\epsilon\widetilde\omega(x)u(x)}]$$

$$= [0, -gAd_{g^{-1}\bar g}\widetilde\xi(\bar x)(\widetilde\omega(x)u(x))]$$

and

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} U(\phi_{\bar\xi_Q}(\epsilon)) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} [u(x), -e^{\epsilon Ad_{\bar g}\widetilde\xi(\bar x)}g(\widetilde\omega(x)u(x))]$$

$$= [0, -gAd_{g^{-1}\bar g}\widetilde\xi(\bar x)(\widetilde\omega(x)u(x))].$$

Therefore

$$[\,U, \bar{\xi}_Q\,] = 0. \tag{5.3.35}$$

Substituting (5.3.33)-(5.3.35) into (5.3.32) and evaluating it at $r = \bar{r}$, one gets (5.3.31).

∎

Finally the last term in (5.2.25) can be given as follows. First, at any point $r = (x, g)$ in Q,

$$
\begin{aligned}
D_r \xi(r) \cdot u_r &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \xi(r + \epsilon u_r) \\
&= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} g \cdot e^{-\epsilon \widetilde{\omega}(x) u_x} \cdot \widetilde{\xi}(x + \epsilon u_x) \cdot e^{\epsilon \widetilde{\omega}(x) u_x} g^{-1} \\
&= g \cdot (-(\widetilde{\omega}(x) u_x)\widetilde{\xi}(x) + \widetilde{\xi}(x)(\widetilde{\omega}(x) u_x) + D_x \widetilde{\xi}(x) \cdot u_x) \cdot g^{-1} \\
&= g \cdot ([\,\widetilde{\xi}(x), \widetilde{\omega}(x) u_x\,] + D_x \widetilde{\xi}(x) \cdot u_x) \cdot g^{-1}
\end{aligned}
$$

which is a $\mathcal{G}$-valued vector. Then

$$[D\xi(r) \cdot u_r]_Q(r) = [0, g \cdot ([\,\widetilde{\xi}(x), \widetilde{\omega}(x) u_x\,] + D_x \widetilde{\xi}(x) \cdot u_x)].$$

Therefore,

$$
\begin{aligned}
\Xi(r)(u_r, v_r) &= K(x, g)([v_x, -g(\widetilde{\omega}(x) v_x)], [0, g \cdot ([\,\widetilde{\xi}(x), \widetilde{\omega}(x) u_x\,] + D_x \widetilde{\xi}(x) \cdot u_x)]) \\
&= K(x, e)([v_x, -(\widetilde{\omega}(x) v_x)], [0, [\,\widetilde{\xi}(x), \widetilde{\omega}(x) u_x\,] + D_x \widetilde{\xi}(x) \cdot u_x]) \\
&\triangleq \widetilde{\Xi}(x)(u_x, v_x). \tag{5.3.36}
\end{aligned}
$$

By collecting (5.3.17), (5.3.24), (5.3.25), (5.3.31) and (5.3.36), Theorem 5.2.5 leads to the following result.

Theorem 5.3.7: If $q(\cdot) = (x(\cdot), g(\cdot))$ is a solution of constrained dynamic equations (5.2.5) on principal fiber bundle $(B \times G, B, \pi, G)$ with connection in (5.3.2), then its projection $x(\cdot)$ in B satisfies unconstraint equation

$$
\begin{aligned}
\frac{d}{dt} D_2 \widetilde{L}^\xi(x, v_x) \cdot u_x - D_1 \widetilde{L}^\xi(x, v_x) \cdot u_x = {}& \widetilde{F}(x, v_x) \cdot u_x + d\widetilde{\omega}_\xi(x)(u_x, v_x) \\
& + \Gamma(x)(v_x, u_x, v_x) + \langle \widetilde{\mathbb{II}}(x)\widetilde{\xi}(x), \widetilde{\Omega}(x)(u_x, v_x) \rangle \\
& + \langle (D_x \widetilde{\mathbb{II}}(x) \cdot u_x)\widetilde{\xi}(x), \widetilde{\xi}(x) \rangle \\
& + \langle \widetilde{\mathbb{II}}(x)\widetilde{\xi}(x), D_x \widetilde{\xi}(x) \cdot u_x \rangle \\
& + \langle \widetilde{\mathbb{II}}(x)\widetilde{\xi}(x), [\,\widetilde{\omega}(x) u_x, \widetilde{\xi}(x)\,] \rangle \\
& - \widetilde{\Xi}(x)(u_x, v_x) \tag{5.3.37}
\end{aligned}
$$

137

for any $u_x \in T_x B$, where

$$\widetilde{L}^\xi(x, v_x) \triangleq \frac{1}{2} \widetilde{K}(x)(v_x, v_x) - (\widetilde{V}(x) + \frac{1}{2} \langle \widetilde{\mathbb{I}}(x)\widetilde{\xi}(x), \widetilde{\xi}(x) \rangle)$$

$$\widetilde{\Omega}(x)(u_x, v_x) \triangleq d\widetilde{\omega}(x)(u_x, v_x) - [\, \widetilde{\omega}(x) \cdot u_x, \widetilde{\omega}(x) \cdot v_x \,]$$

$$\Gamma(x)(v_x, u_x, v_x) \triangleq K(x, e)([v_x, -\widetilde{\omega}(x) \cdot v_x], [0, \widetilde{\Omega}(x)(u_x, v_x)]),$$

$$\widetilde{\Xi}(x)(u_x, v_x) \triangleq K(x, e)([v_x, -(\widetilde{\omega}(x)v_x)], [0, [\, \widetilde{\xi}(x), \widetilde{\omega}(x)u_x \,] + D_x\widetilde{\xi}(x) \cdot u_x])$$

and $\widetilde{F}$ and $\widetilde{\omega}_\xi$ are defined in (5.3.24) and (5.3.26), respectively. ∎

Remarks 5.3.8:

(1) If we define a function

$$\widetilde{L}(x, v_x) \triangleq \frac{1}{2} \widetilde{K}(x)(v_x, v_x) - \widetilde{V}(x),$$

by considering (5.3.16), the dynamic equation (5.3.37) can be replace by

$$
\begin{aligned}
\frac{d}{dt} D_2 \widetilde{L}(x, v_x) \cdot u_x - D_1 \widetilde{L}(x, v_x) \cdot u_x = & \widetilde{F}(x, v_x) \cdot u_x + d\widetilde{\omega}_\xi(x)(u_x, v_x) \\
& + \Gamma(x)(v_x, u_x, v_x) + \langle \widetilde{\mathbb{I}}(x)\widetilde{\xi}(x), \widetilde{\Omega}(x)(u_x, v_x) \rangle \\
& + \frac{1}{2} \langle (D_x\widetilde{\mathbb{I}}(x) \cdot u_x)\widetilde{\xi}(x), \widetilde{\xi}(x) \rangle \\
& + \langle \widetilde{\mathbb{I}}(x)\widetilde{\xi}(x), [\, \widetilde{\omega}(x)u_x, \widetilde{\xi}(x) \,] \rangle \\
& - \widetilde{\Xi}(x)(u_x, v_x).
\end{aligned}
\tag{5.3.38}
$$

(2) If $\widetilde{\xi}(x) \equiv 0$, Equation (5.3.37) becomes

$$\frac{d}{dt} D_2 \widetilde{L}(x, v_x) \cdot u_x - D_1 \widetilde{L}(x, v_x) \cdot u_x = \widetilde{F}(x, v_x) \cdot u_x + \Gamma(x)(v_x, u_x, v_x), \tag{5.3.39}$$

which is the equation derived in [20].

(3) As long as $x(\cdot)$ is determined from (5.3.37) or (5.3.38), its horizontal lift $r(\cdot)$ can be found from solving (5.3.10), and the solution of original constrained equation (5.2.5) can be found by using the formula in Remark 5.2.6(3). This procedure is called *reconstruction* in [26].

(4) If one assumes G acts on Q on the *right* in the definition of principal fiber bundle and simple mechanical system with symmetry and the mapping $\xi : Q \to \mathcal{G}'$ defined in Constraint Hypothesis 5.2.1 is G-equivariant with respect to *right* action, i.e.,

$\xi(q \cdot g) = Ad_{g^{-1}}\xi(q)$, then to obtain the reduced dynamic equation, one only needs to change signs in front of all the Lie brackets in (5.3.37) or (5.3.38).

(5) One can check that if G is Abelian, (5.3.37) or (5.3.38) specializes to (5.1.9). ∎

5.4 Lagrangian Reduction

In Section 5.2, we formulated constrained Lagrangian dynamics on horizontal distribution for any principal connection, by which the nonholonomic constraints are constructed (cf. Constraint Hypothesis 5.2.1). Here, we consider a special connection, namely, *mechanical connection* which has been considered in Section 4.1.

Recall that, given a simple mechanical system with symmetry (Q, K, V, G) with its Lagrangian given in (5.2.2), the lift of the G-action to TQ induces an equivariant momentum map $\mathbf{J} : TQ \to \mathcal{G}^*$ given by

$$\langle \mathbf{J}(q, v_q), \zeta \rangle = K(q)(v_q, \zeta_Q(q)), \quad \forall \zeta \in \mathcal{G}. \tag{5.4.1}$$

We assume that the exterior force acting on the system satisfies (5.2.6) and leaves the momentum map invariant, i.e., the motion of the system $q(\cdot) = \{q(t), t \geq 0\}$ satisfies $\mathbf{J}(q, \dot{q}) = \mu$ for a constant $\mu \in \mathcal{G}^*$. If μ is a regular value of $\mathbf{J}$, the interesting question is what the dynamics on subspace $\mathcal{S} = \mathbf{J}^{-1}(\mu) \subset TQ$ is. In Section 5.1, we have considered the same problem by treating the conserved momentum map as the constraints of the system with Abelian symmetry. Here, although the symmetry group is non-Abelian in general, we will follow the same idea as we had in Section 5.1, but using the geometric method we developed in Sections 5.2 and 5.3.

Consider a principal fiber bundle given by

$$\wp = (Q, B, \pi, G_\mu), \tag{5.4.2}$$

where $G_\mu = \{g \in G \mid Ad^*_{g^{-1}}\mu = \mu\}$, for constant $\mu \in \mathcal{G}^*$ given above, is an isotropy group. Let $\mathcal{G}_\mu$ be the Lie algebra of G_μ. As we have shown in Section 4.1.2, on this bundle, the mechanical connection is given by a $\mathcal{G}_\mu$-valued one-form:

$$\omega(q) : T_qQ \to \mathcal{G}_\mu$$
$$v_q \mapsto \mathbb{I}_\mu^{-1}(q)\mathbf{J}(q, v_q), \quad \forall q \in Q, \tag{5.4.3}$$

139

where $\mathbb{I}_\mu(q) : \mathcal{G}_\mu \to \mathcal{G}_\mu^*$ is called μ-locked inertia tensor, which is defined in the same way as we did in (4.1.9), but restricted on $\mathcal{G}_\mu$. Now the subspace $\mathcal{S} = \mathbf{J}^{-1}(\mu)$ can be represented as

$$\mathcal{S} = \{ (q, v_q) \in TQ \mid \omega(q)(v_q) = \mathbb{I}_\mu^{-1}(q)\mu \}, \tag{5.4.4}$$

which is of the same form as (5.2.4) in Constraint Hypothesis 5.2.1. Indeed, for any $g \in \mathcal{G}_\mu$,

$$\mathbb{I}_\mu^{-1}(g \cdot q)\mu = \mathbb{I}_\mu^{-1}(g \cdot q)Ad_g^*\mu = \mathbb{I}_\mu^{-1}(g \cdot q)Ad_g^*\mathbf{J}(q, v_q)$$

$$= \mathbb{I}_\mu^{-1}(g \cdot q)\mathbf{J}(T_q\Phi_g \cdot (q, v_q)) = \omega(g \cdot q)(T_q\Phi_g \cdot v_q)$$

$$= Ad_g\omega(q)(v_q) = Ad_g\,\mathbb{I}_\mu^{-1}(q)\mu.$$

With the above setting, formulation of dynamic equation on $\mathbf{J}^{-1}(\mu)$ can be given by (5.2.5) with $\xi(q) = \mathbb{I}_\mu^{-1}(q)\mu$. Following Theorem 5.2.4, we can also write down the dynamics of the constrained system on the horizontal distribution. Since the connection now is a special one, Equation (5.2.25) can be simplified. We first show the following results.

Proposition 5.4.1: In Theorem 5.2.5, if $\xi(q) = \mathbb{I}_\mu^{-1}(q)\mu$,

(1) the function $V^\xi(q)$ is the amended potential function $V_\mu(q)$;

(2) the one-form ω_ξ is the μ-component of mechanical connection form;

(3) For mechanical connection,

$$\Xi(r)(\cdot, \cdot) \equiv 0.$$

Proof: (1) Indeed,

$$V^\xi(q) = V(q) + \frac{1}{2}\,K(q)([\mathbb{I}_\mu^{-1}(q)\mu]_Q(q), [\mathbb{I}_\mu^{-1}(q)\mu]_Q(q))$$

$$= V(q) + \frac{1}{2}\,\langle\mathbb{I}_\mu(q) \cdot \mathbb{I}_\mu^{-1}(q)\mu, \mathbb{I}_\mu^{-1}(q)\mu\rangle$$

$$= V(q) + \frac{1}{2}\,\langle\mu, \mathbb{I}_\mu^{-1}(q)\mu\rangle$$

$$= V_\mu(q).$$

(2) The μ-component of a k-form $\alpha^k \in \varpi^k(Q; \mathcal{G}_\mu)$ is a real-valued k-form on Q, denoted by α_μ^k, and is defined by

$$\langle\alpha_\mu^k(q), v_q\rangle = \langle\mu, \alpha^k(q)(v_q)\rangle. \tag{5.4.5}$$

140

From the definition of mechanical connection given in (5.4.3), we have

$$\langle \mu, \omega(q)(v_q) \rangle = \langle \mu, \mathbb{I}_\mu^{-1}(q)\mathbf{J}(q, v_q) \rangle$$

$$= \langle \mathbb{I}_\mu^{-1}(q)\mu, \mathbf{J}(q, v_q) \rangle$$

$$= K(q)((\mathbb{I}_\mu^{-1}(q)\mu)_Q(q), v_q)$$

$$= \langle K^\flat(q)((\mathbb{I}_\mu^{-1}(q)\mu)_Q(q)), v_q \rangle.$$

Therefore, from (5.4.5), the μ-component of the mechanical connection form is

$$\omega_\mu(q) = K^\flat(q)([\mathbb{I}_\mu^{-1}(q)\mu]_Q(q)). \tag{5.4.6}$$

From the condition of this Proposition and the expression of ω_ξ, (2) is proved.

(3) Since horizontal-vertical splitting of mechanical connection is an orthogonal splitting with respect to the Riemannian metric K, the claim follows from the definition of $\Xi(r)(\cdot, \cdot)$. ∎

After identifying ω_ξ as μ-component of mechanical connection form ω_μ, we have the following properties.

Proposition 5.4.2:

(1) For a fixed point $\bar{q} \in Q$ and any vector field $Y \in \mathfrak{X}(Q)$,

$$d\omega_\mu(\bar{q})([\xi(\bar{q})]_Q(\bar{q}), Y(\bar{q})) \equiv 0; \tag{5.4.7}$$

(2) $d\omega_\mu$ is the μ-component of the curvature form, Ω_μ, of the mechanical connection.

Proof:

(1) First, we extend the tangent vector $[\xi(\bar{q})]_Q(\bar{q})$ to a fundamental vector field (at $\bar{q}$) $\bar{\xi}_Q(q)$, where $\bar{\xi} = \xi(\bar{q})$. Then we have formula

$$d\omega_\mu(\bar{\xi}_Q, Y) = \bar{\xi}_Q \cdot \omega_\mu(Y) - Y \cdot \omega_\mu(\bar{\xi}_Q) - \omega_\mu([\bar{\xi}_Q, Y])$$

$$= \langle \mu, \bar{\xi}_Q \cdot \omega(Y) \rangle - \langle \mu, Y \cdot \omega(\bar{\xi}_Q) \rangle - \langle \mu, \omega([\bar{\xi}_Q, Y]) \rangle. \tag{5.4.8}$$

The second term in (5.4.8) is always zero since $\omega(\bar{\xi}_Q) = \bar{\xi}$ is a constant in $\mathcal{G}_\mu$. If Y is horizontal, so is $[\bar{\xi}_Q, Y]$ (cf. [33]). This implies $\omega(Y) = \omega([\bar{\xi}_Q, Y]) = 0$, and consequently $d\omega_\mu(\bar{\xi}_Q, Y) = 0$. If Y is fundamental, say $Y = \zeta_Q$ for some $\zeta \in \mathcal{G}_\mu$, the

first term of (5.4.8) is zero since $\omega(Y)$ is a constant. In addition,

$$
\begin{aligned}
\langle \mu, \omega([\, \xi_Q, Y \,]) \rangle &= -\langle \mu, \omega([\, \xi, \zeta \,]_Q) \rangle \\
&= -\langle \mu, [\, \xi, \zeta \,] \rangle \\
&= -\frac{d}{dt}\Big|_{t=0} \langle \mu, Ad_{exp(t\xi)}\zeta \rangle \\
&= -\frac{d}{dt}\Big|_{t=0} \langle Ad^*_{exp(t\xi)}\mu, \zeta \rangle \\
&= -\frac{d}{dt}\Big|_{t=0} \langle \mu, \zeta \rangle = 0.
\end{aligned}
$$

This completes the proof of (1).

(2) From the structural equation in (4.1.4), we have

$$
\begin{aligned}
\Omega_\mu(X, Y) &= \langle \mu, \Omega(X, Y) \rangle \\
&= \langle \mu, [\, \omega(X), \omega(Y) \,] \rangle + \langle \mu, d\omega(X, Y) \rangle.
\end{aligned}
$$

But $\langle \mu, [\, \omega(X), \omega(Y) \,] \rangle = 0$ by the proof of (1) above. This proves the result. ∎

Now using the above observations, we can restate the Theorem 5.2.5 for the systems with mechanical connection as follows.

Theorem 5.4.3: If $q(\cdot)$ is the motion of a Lagrangian system with symmetry and preserves the equivariant momentum map $\mathbf{J}(q, v_q) = \mu$, then its restriction, $r(\cdot)$, on the horizontal distribution determined by the mechanical connection satisfies the dynamic equation

$$
\frac{d}{dt}D_2 L_\mu(r, v_r)\cdot u_r - D_1 L_\mu(r, v_r)\cdot u_r = F(r, v_r + [\mathbb{I}_\mu^{-1}(r)]_Q(r))\cdot u_r + \Omega_\mu(r)(u_r, v_r), \quad (5.4.9)
$$

where

$$
L_\mu(r, v_r) = \frac{1}{2}K(r)(v_r, v_r) - (V(r) + \frac{1}{2}\langle \mu, \mathbb{I}_\mu^{-1}(r)\mu \rangle) \qquad (5.4.10)
$$

and Ω_μ is μ-component of the curvature form of the mechanical connection. ∎

Having dynamic equation on the horizontal distribution, we intend to drop the dynamics of the system down to the base space Q/G_μ. To this end, we first define a lift mapping from curves in base space to curves in total space. Let $q(\cdot) = \{q(t), t \geq 0\}$ be a curve in Q and $x(\cdot)$ be the bundle projection of $q(\cdot)$, i.e., $x(\cdot) = \pi(q(\cdot))$. Define a mapping (cross-section)

$$
\rho : Q/G_\mu \to Q \qquad (5.4.11)
$$

142

such that $r(\cdot) \triangleq \rho(x(\cdot))$ is a horizontal curve in Q, i.e., at each $t = \bar{t}$,

$$v_r(\bar{t}) = \mathcal{H}(\rho(x(\bar{t}))) \cdot \dot{x}(\bar{t}) \triangleq T_{x(\bar{t})}\rho(x(\bar{t})) \cdot \dot{x}(\bar{t}) \in \mathbf{H}_{\rho(x(\bar{t}))}. \qquad (5.4.12)$$

If the principal bundle is a trivial one, the mapping $\mathcal{H}(r)$ has been given explicitly in Remark 5.3.1. In particular, one can show that now the pull-down connection form $\widetilde{\omega}(x)$ is of the form: $\widetilde{\mathbb{I}}^{-1}(x)\jmath(x)$, where the map $\jmath(x)$ has been defined in Subsection 4.1.2.

With the above defined mappings, one can define the induced Riemannian metric, potential function and μ-locked inertia tensor on the base space $B = Q/G_\mu$ as follows.

$$\widetilde{K}(x)(v_x, w_x) \triangleq K(\rho(x))(\mathcal{H}(\rho(x)) \cdot v_x, \mathcal{H}(\rho(x)) \cdot w_x) \quad \forall v_x, w_x \in T_x B,$$
$$(5.4.13)$$

$$\widetilde{V}(x) \triangleq (V \circ \rho)(x), \qquad (5.4.14)$$

$$\widetilde{\mathbb{I}}_\mu(x) \triangleq (\mathbb{I}_\mu \circ \rho)(x). \qquad (5.4.15)$$

Remark 5.4.4: The above objects are well defined since they all are G_μ-invariant. ∎

With these newly defined objects in B, the Lagrangian L_μ in (5.4.10) can also be induced on B:

$$
\begin{aligned}
L_\mu(r, v_r) &= L_\mu(\rho(x), \mathcal{H}(\rho(x)) \cdot v_x) \\
&= \frac{1}{2} K(\rho(x))(\mathcal{H}(\rho(x)) \cdot v_x, \mathcal{H}(\rho(x)) \cdot v_x) \\
&\quad - (V(\rho(x)) + \frac{1}{2}\langle \mu, \mathbb{I}_\mu^{-1}(\rho(x))\mu \rangle) \\
&= \frac{1}{2}\widetilde{K}(x)(v_x, v_x) - (\widetilde{V}(x) + \frac{1}{2}\langle \mu, \widetilde{\mathbb{I}}_\mu^{-1}(x)\mu \rangle) \\
&\triangleq \widetilde{L}_\mu(x, v_x), \qquad (5.4.16)
\end{aligned}
$$

which is a function defined on $T(Q/G_\mu)$.

Now we are ready to express (5.4.9) on the base space Q/G_μ. Let $u_r = \mathcal{H}(\rho(x)) \cdot u_x$ for any $u_x \in T_x B$. Then, the first term on the left-hand-side of (5.4.9) is

$$
\begin{aligned}
\frac{d}{dt} D_2 L_\mu(r(t), v_r(t)) \cdot u_r &= \frac{d}{dt} K(r(t))(v_r(t), u_r) \\
&= \frac{d}{d\tau}\bigg|_{\tau=t} K(\rho(x(\tau)))(\mathcal{H}(\rho(x(\tau))) \cdot v_x, \mathcal{H}(\rho(x(t))) \cdot u_x) \\
&= \frac{d}{dt}\widetilde{K}(x(t))(v_x(t), u_x) - K(\rho(x))(\mathcal{H}(\rho(x)) \cdot v_x, \frac{d}{dt}\mathcal{H}(\rho(x(t))) \cdot u_x) \\
&= \frac{d}{dt} D_2 \widetilde{L}_\mu(x, v_x) \cdot u_x - [K(r)(v_r, (D_q\mathcal{H}(r) \cdot v_r) \cdot u_x)]|_{r, v_r}, \qquad (5.4.17)
\end{aligned}
$$

143

where $[\,\cdot\,]|_{r,v_r} \triangleq [\,\cdot\,]|_{r=\rho(x),v_r=\mathcal{H}(\rho(x))\cdot v_x}$. The second term of (5.4.9) can be expressed explicitly as

$$D_r L_\mu(r, v_r) \cdot u_r = \frac{1}{2}(D_r K(r) \cdot u_r)(v_r, v_r) - D_r V(r) \cdot u_r - \frac{1}{2}\langle \mu, (D_r \mathbb{I}_\mu^{-1}(r) \cdot u_r)\mu\rangle.$$

Here

$$
\begin{aligned}
(D_r K(r) \cdot u_r)(v_r, v_r) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} K(r + \epsilon u_r)(v_r, v_r) \\
&= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} K(\rho(x + \epsilon u_x))(\mathcal{H}(\rho(x)) \cdot v_x, \mathcal{H}(\rho(x)) \cdot v_x) \\
&= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} [\widetilde{K}(x + \epsilon u_x)(v_x, v_x) \\
&\qquad - K(\rho(x))(\mathcal{H}(\rho(x + \epsilon u_x)) \cdot v_r, \mathcal{H}(\rho(x + \epsilon u_x)) \cdot v_r)] \\
&= (D_x \widetilde{K}(x) \cdot u_x)(v_x, v_x) \\
&\qquad - 2[K(r)(v_r, (D_q \mathcal{H}(r) \cdot u_r) \cdot v_x)]|_{r,u_r,v_r},
\end{aligned}
$$

where $[\,\cdot\,]|_{r,u_r,v_r} \triangleq [\,\cdot\,]|_{r=\rho(x),u_r=\mathcal{H}(\rho(x))\cdot u_x,v_r=\mathcal{H}(\rho(x))\cdot v_x}$. and, following the same approach, we have

$$D_r V(r) \cdot u_r = D_x \widetilde{V}(x) \cdot u_x$$

$$D_r \mathbb{I}_\mu(r)^{-1} \cdot u_r = D_x \widetilde{\mathbb{I}}_\mu(x)^{-1} \cdot u_x.$$

Then, the second term of (5.4.9) now is of the form

$$D_r L_\mu(r, v_r) \cdot u_r = D_1 \widetilde{L}_\mu(x, v_x) \cdot u_x - [K(r)(v_r, (D_q \mathcal{H}(r) \cdot u_r) \cdot v_x)]|_{r,u_r,v_r}. \quad (5.4.18)$$

From (5.4.17) and (5.4.18), the left-hand-side of (5.4.9) becomes

$$
\begin{aligned}
\frac{d}{dt} D_2 L_\mu(r(t), v_r(t)) \cdot u_r &- D_r L_\mu(r, v_r) \cdot u_r = \\
&\frac{d}{dt} D_2 \widetilde{L}_\mu(x, v_x) \cdot u_x - D_1 \widetilde{L}_\mu(x, v_x) \cdot u_x + \Gamma(x)(v_x, u_x, v_x),
\end{aligned}
\quad (5.4.19)
$$

where

$$\Gamma(x)(v_x, u_x, v_x) = [K(r)(v_r, (D_q \mathcal{H}(r) \cdot u_r) \cdot v_x - (D_q \mathcal{H}(r) \cdot v_r) \cdot u_x)]|_{r,u_r,v_r}.$$

From the previous section, we have known that if the principal fiber bundle is a trivial bundle,

$$[(D_q \mathcal{H}(r) \cdot u_r) \cdot v_x - (D_q \mathcal{H}(r) \cdot v_r) \cdot u_x)]|_{r,u_r,v_r} = Ver([\mathcal{H}(\rho(x)) \cdot u_x, \mathcal{H}(\rho(x)) \cdot v_x]).$$

144

Note that a principal bundle is locally trivial. Thus, the above result is also true in general. Since, for mechanical connection, the horizontal-vertical splitting is an orthogonal one with respect to the Riemannian metric K, immediately, we have

$$\Gamma(x)(v_x, u_x, v_x) \equiv 0. \tag{5.4.20}$$

Return to the right-hand-side of (5.4.9). Recall that Ω_μ is the μ-component of curvature form of the mechanical connection defined by

$$\Omega_\mu(q)(v_q, w_q) = \langle \mu, \Omega(q)(v_q, w_q) \rangle, \quad \forall v_q, w_q \in T_q Q.$$

Since Ω is G_μ-equivariant, Ω_μ is G-invariant. Therefore, using the lift mappings in (5.4.11) and (5.4.12), we can define a two-form on base space B by

$$\widetilde{\Omega}_\mu(x)(u_x, v_x) \triangleq \Omega_\mu(\rho(x))(\mathcal{H}(\rho(x)) \cdot u_x, \mathcal{H}(\rho(x)) \cdot v_x). \tag{5.4.21}$$

Finally, using the G_μ-equivariant property of the exterior force (cf. (5.2.6)), we can define exterior force on TB by

$$\widetilde{F}(x, v_x) \cdot u_x \triangleq F(\rho(x), \mathcal{H}(\rho(x)) \cdot v_x + [\widetilde{\mathbb{I}}_\mu^{-1}(x)\mu]_Q(\rho(x))) \cdot (\mathcal{H}(\rho(x)) \cdot u_x). \quad \forall u_x \in T_x B \tag{5.4.22}$$

From (5.4.19) - (5.4.22), we have the following theorem.

Theorem 5.4.5: If $q(\cdot)$ is a motion of Lagrangian system with symmetry and keeps momentum map conserved, then its projection $x(\cdot)$ in Q/G_μ satisfies dynamic equation

$$\frac{d}{dt} D_2 \widetilde{L}_\mu(x, v_x) \cdot u_x - D_1 \widetilde{L}_\mu(x, v_x) \cdot u_x = \widetilde{F}(x, v_x) \cdot u_x + \widetilde{\Omega}_\mu(x)(u_x, v_x)$$

for any $u_x \in T_x(Q/G_\mu)$, where

$$\widetilde{L}_\mu(x, v_x) = \frac{1}{2} \widetilde{K}(x)(v_x, v_x) - (\widetilde{V}(x) + \frac{1}{2}\langle \mu, \widetilde{\mathbb{I}}_\mu^{-1}(x)\mu \rangle),$$

$\widetilde{F}$ and $\widetilde{\Omega}_\mu$ are defined in (5.4.22) and (5.4.21), respectively. ∎

Remarks 5.4.6:

(1) In the literature of the analytic mechanics, the reduction theory for Hamiltonian systems has been well developed [26,28,29] and has been successfully applied to many problems in engineering. However, since in many physical problems,

Lagrangian dynamics is the natural starting point, the construction of a reduction theory directly applicable to Lagrangian systems becomes a challenge. Recent work done by Marsden and Scheurle [27], and Bloch and Crouch [7] have contributed to this goal. In [27], the reduction is realized by including conservative gyroscopic (magnetic) force into the variational principle in the sense of Lagrange and d'Alembert. The main advantage of this approach is that the reduction procedure is directly comparable with the one in Hamiltonian case.

(2) Once the motion of the system in Q/G_μ is found, the motion of the system in Q can be determined using a *reconstruction procedure*, cf. [26].

(3) Lagrangian reduction can be applied to study the stability of relative equilibria of various mechanical systems. For example, also see [27]. ∎

5.5 Examples

To illustrate the results of this chapter, we apply the theorems in this chapter to some representative physical systems.

Example 5.5.1: *Reduced dynamic equation for planar 3-body system.*

In [34,43,44], the authors investigate the dynamics and control of planar multibody systems. In particular, the relative equilibria and their stability are investigated using the energy-Casimir method. Bifurcation phenomena are explored by standard methods in Hamiltonian mechanics. Here, following [27], we shall use Lagrangian reduced dynamics to compute the relative equilibria and then discover bifurcations via the movement of the eigenvalues of the linearized dynamics at *fundamental* relative equilibria.

Consider a planar three-body system shown in Figure 5.5.1. The configuration space now is $S^1 \times S^1 \times S^1$ and the symmetry group is S^1. Since S^1 is Abelian, we only need to use the formula in Section 5.1 for reduction. The Lagrangian of this system, after ignoring the potential, can be written as

$$L(\phi, \dot{\phi}, \dot{\theta}_3) = \frac{1}{2}\dot{\phi}^T K_{11}(\phi)\dot{\phi} + \dot{\phi}^T K_{12}(\phi)\dot{\theta}_3 + \frac{1}{2}K_{22}(\phi)(\dot{\theta}_3)^2, \qquad (5.5.1)$$

where $\phi = (\phi_1, \phi_2)^T$. For the sake of simplicity, we assume the planar moments $I_i = 1$ and the masses $m_i = 3$, for $i = 1, 2, 3$. And, $d_{21} = d_{23} = d_{32} = 1$, but $d_{12} = l > 0$
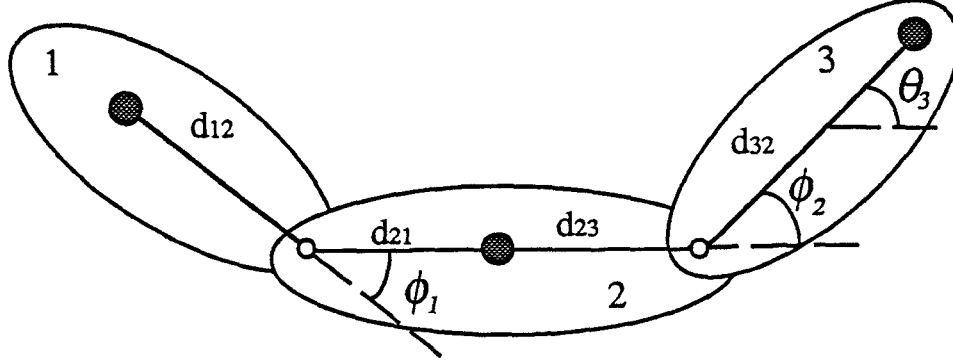
Figure 5.5.1 A planar three-body systems

which is the only variable parameter for the system. Then, the matrices in (5.1.18) are of the form

$$K_{11}(\phi) = \begin{pmatrix} 6l^2 + 1 & 9lcos(\phi_1) + 6l^2 + 1 \\ 9lcos(\phi_1) + 6l^2 + 1 & 18lcos(\phi_1) + 6l^2 + 20 \end{pmatrix}$$

$$K_{12}(\phi) = \begin{pmatrix} 3lcos(\phi_2 + \phi_1) + 9lcos(\phi_1) + 6l^2 + 1 \\ 3lcos(\phi_2 + \phi_1) + 9cos(\phi_2) + 18lcos(\phi_1) + 6l^2 + 20 \end{pmatrix}$$

$$K_{22}(\phi) = 6lcos(\phi_2 + \phi_1) + 18cos(\phi_2) + 18lcos(\phi_1) + 6l^2 + 27.$$

Using the above matrices the reduced Lagrangian, μ-component of the curvature form and, consequently, the reduced dynamics can be easily determined from (5.1.14)-(5.1.16). In particular, from (5.1.14), the equilibria of the reduced system are determined by the critical points of K_{22} which is just the locked inertia of the system. These equilibria are the relative equilibria of unreduced system. It is readily seen that there are at least four critical points for K_{22}, that is, $\phi = \{(0,0),(0,\pi),(\pi,0),(\pi,\pi)\}$, which are called *fundamental* relative equilibria [43]. The stability of these equilibria can be easily determined by looking at the spectrum of linearized equation of reduced dynamics. In addition, we can also see the movement of the spectrum in response to change of certain parameters (e.g., l) and the bifurcation points. In the following, we study the linearized equation at (π,π) only.

One can show that the linearized equation at (π, π) of the reduced dynamics is

$$M\ddot{\phi} + \Lambda\phi = 0, \tag{5.5.2}$$

where

$$M = \begin{pmatrix} 24\,l^4 - 48\,l^3 + 52\,l^2 - 32\,l + 24 & 12\,l^4 - 48\,l^3 + 62\,l^2 - 28\,l - 6 \\ 12\,l^4 - 48\,l^3 + 62\,l^2 - 28\,l - 6 & 66\,l^4 - 276\,l^3 + 505\,l^2 - 452\,l + 177 \end{pmatrix}$$

and

$$\Lambda = \begin{pmatrix} -2\,l\,\mu^2 & l\,\mu^2 \\ l\,\mu^2 & l\,\mu^2 - 3\,\mu^2 \end{pmatrix}.$$

Then, the eigenvalues of this system are the roots of the equation

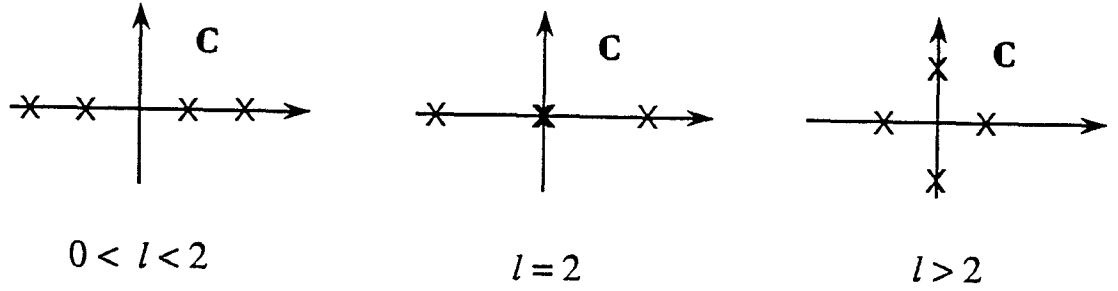$$det(\lambda^2 M + \Lambda) = 0.$$

Figure 5.5.2 Eigenvalues of linearized equation

Figure 5.1.2 shows the positions of four eigenvalues in the complex plane for three different values of l. Since the eigenvalues of the linearized system (5.1.19) are also the eigenvalues of Hamiltonian system

$$\dot{q} = M^{-1}p$$

$$\dot{p} = -\Lambda q$$

for $p = M\dot{q}$, following the analysis in [34], bifurcation occurs at $l = 2$ necessarily. In fact, one can see this bifurcation from the contours of amended potential function shown in Figure 5.1.3. ∎
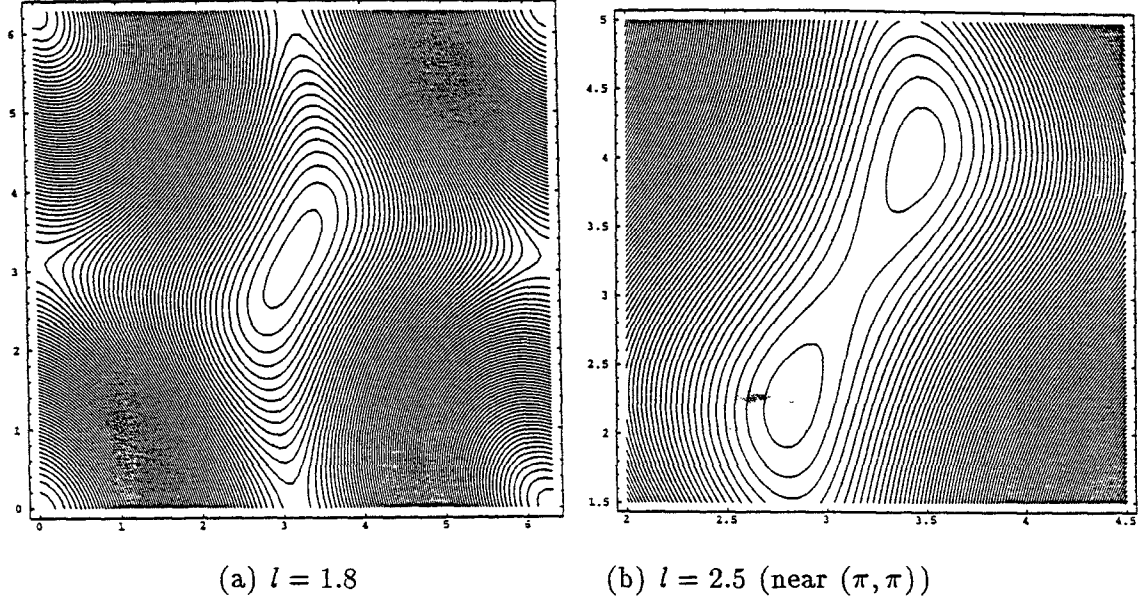
(a) $l = 1.8$ (b) $l = 2.5$ (near (π, π))

Figure 5.5.3 Contours of amended potential of planar 3-body system

Example 5.5.2: *Rolling a homogeneous sphere on a rotating horizontal plane* [32].

Let 0-XYZ be a coordinates system fixed in inertial space. Let a homogeneous sphere roll without sliding on a plane (or platform) which rotates about the Z axis with a constant angular velocity C, see Figure 5.5.4.

Since the sphere will not move along the Z direction, its configuration space is $\Re^2 \times SO(3)$ which will be parametrized by (p, A), where $p = (x, y)$ gives the location of center of the sphere, or the contact point of the sphere and the plane, A gives the orientation of the sphere relative to inertial space or coordinates system 0-XYZ. Let $\zeta = (\omega_x, \omega_y, \omega_z)^T$ be the vector of angular velocity of the sphere in the inertial frame. Then the Lagrangian of this system is

$$
\begin{aligned}
L &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mk^2(\omega_x^2 + \omega_y^2 + \omega_z^2) \\
&= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mk^2 \langle \widehat{\zeta}A, \widehat{\zeta}A \rangle.
\end{aligned}
\tag{5.5.3}
$$

The rolling-without-sliding constraint on the sphere is given by two equations:

$$
\begin{cases}
\dot{x} - a\omega_y + Cy = 0 \\
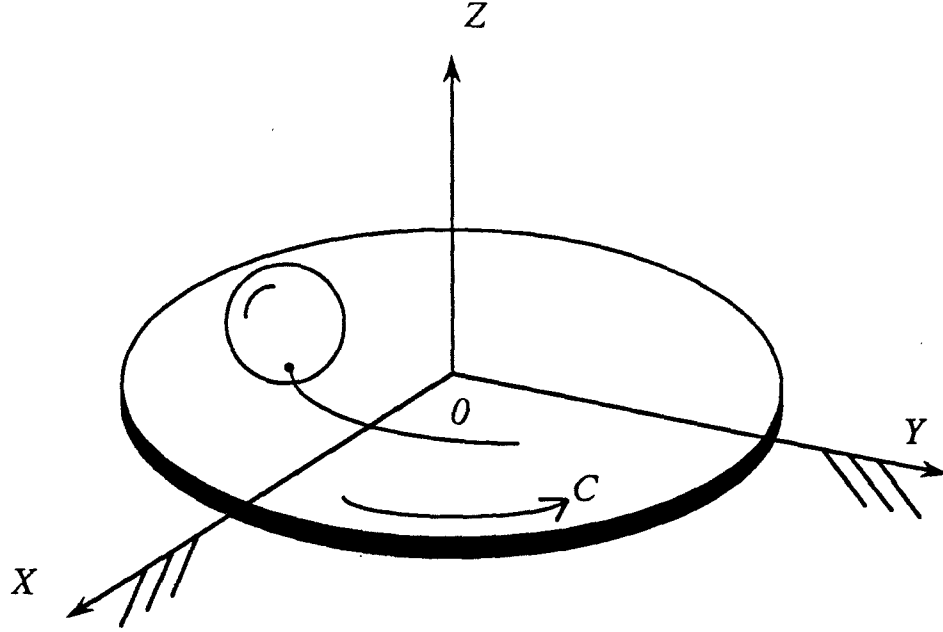\dot{y} + a\omega_x - Cx = 0.
\end{cases}
$$

149

Figure 5.5.4 Rolling a homogeneous sphere on a rotating plane.

We also assume that there is an exterior force, denoted by $F = (F_x, F_y)$, acting at the center of the sphere and along any direction perpendicular to Z axis, where F_x and F_y can be any time dependent functions on $T\Re^2$. It is clear that since there is no torque along the Z axis, the angular velocity of the sphere about the Z axis is a constant, denoted by c. As we did in the second part of Example 2.2.24, we assume the conserved ω_z as an *a priori* constraint. Then the constraint equations can be written as

$$\zeta + \widetilde{\omega}(p)v_p = \widetilde{\xi}(p), \tag{5.5.4}$$

where $v_p = (\dot{x}, \dot{y})^T$,

$$\widetilde{\omega}(p) = \begin{pmatrix} 0 & \frac{1}{a} \\ -\frac{1}{a} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{\xi}(p) = \begin{pmatrix} Cx/a \\ Cy/a \\ c \end{pmatrix},$$

or

$$A^T[\zeta + \widetilde{\omega}(p)v_p]\widehat{\ }A = A^T\widehat{\widetilde{\xi}}(p)A. \tag{5.5.5}$$

Define a *right* action of the Lie group $SO(3)$ on Q by

$$\Phi : SO(3) \times Q \to Q$$

$$(B, (x, y, A)) \mapsto (x, y, AB). \tag{5.5.6}$$

150

Then, it is easy to check that the left-hand-side of (5.5.5) defines a connection on the right-principal fiber bundle $\wp = (\Re^2 \times SO(3), \Re^2, \pi, SO(3))$, the right-hand-side of (5.5.5) satisfies the Constraint Hypothesis 5.2.1 and Lagrangian in (5.5.3) is G-invariant. Since all the conditions for reduction in Section 5.2 and 5.3 are satisfied with respect to the *right* action, we can use Equation (5.3.37) or (5.3.38), taking care to change the order of vectors in all Lie brackets, thereby obtaining the reduced dynamics on the base space $\Re^2$. Here, for convenience, we rewrite (5.3.38) with respect to right action as follows: for any $p \in B$,

$$
\begin{aligned}
\frac{d}{dt} D_2 \widetilde{L}(p, v_p) \cdot u_p - D_1 \widetilde{L}(p, v_p) \cdot u_p = &\ \widetilde{F}(p, v_p) \cdot u_p + d\widetilde{\omega}_\xi(p)(u_p, v_p) \\
&+ \Gamma(p)(v_p, u_p, v_p) + \langle \widetilde{\mathbb{I}}(p)\widetilde{\xi}(p), \widetilde{\Omega}(p)(u_p, v_p) \rangle \\
&+ \frac{1}{2} \langle (D_p \widetilde{\mathbb{I}}(p) \cdot u_p)\widetilde{\xi}(p), \widetilde{\xi}(p) \rangle \\
&+ \langle \widetilde{\mathbb{I}}(p)\widetilde{\xi}(p), [\,\widetilde{\xi}(p), \widetilde{\omega}(p)u_p\,] \rangle \\
&- \widetilde{\Xi}(p)(u_p, v_p),
\end{aligned}
\tag{5.5.7}
$$

where

$$
\widetilde{L}(p, v_p) = \frac{1}{2}\, \widetilde{K}(p)(v_p, v_p) - \widetilde{V}(p),
$$

$$
\widetilde{\Omega}(p)(u_p, v_p) = d\widetilde{\omega}(p)(u_p, v_p) + [\,\widetilde{\omega}(p) \cdot u_p, \widetilde{\omega}(p) \cdot v_p\,],
$$

$$
\Gamma(p)(v_p, u_p, v_p) = K(p, e)([v_p, -\widetilde{\omega}(p) \cdot v_p], [0, \widetilde{\Omega}(p)(u_p, v_p)]),
$$

$$
\widetilde{\Xi}(p)(u_p, v_p) = K(p, e)([v_p, -(\widetilde{\omega}(p)v_p)], [0, [\,\widetilde{\omega}(p)u_p, \widetilde{\xi}(p)\,] + D_p\widetilde{\xi}(p) \cdot u_p])
$$

and, without change, $\widetilde{F}$ and $\widetilde{\omega}_\xi$ are defined as in (5.3.24) and (5.3.26), respectively.

Let $\widetilde{M} = Diag(m, m, mk^2, mk^2, mk^2)$. Then $\widetilde{L}$ in (5.5.7) is

$$
\widetilde{L} = \frac{1}{2} v_p^T \widetilde{\omega}(p)^T \widetilde{M} \widetilde{\omega}(p) v_p = \frac{1}{2} m \frac{k^2 + a^2}{a^2} (\dot{x}^2 + \dot{y}^2).
\tag{5.5.8}
$$

The other terms in (5.5.7) can be written as follows. Let $u_p = (u_x, u_y)$. Then

$$
\widetilde{F}(p, v_p) u_p = F_x u_x + F_y u_y.
\tag{5.5.9}
$$

Since in this case $\widetilde{K}$ and $\widetilde{\omega}$ do not depend on p, from (5.3.26), for any $w = (w_x, w_y) \in T_p B$,

$$
\widetilde{\omega}_\xi(p)(w) = \frac{mk^2 C}{a^2}(y w_x - x w_y)
$$

and $\widetilde{\mathbb{I}}(p) = mk^2 I_{3\times3}$, which is a constant matrix. The Lie brackets in (5.5.7) have the following form: letting $\{e_1, e_2, e_3\}$ be the standard basis for $so(3)$,

$$[\,\widetilde{\omega}(p)u_p, \widetilde{\omega}(p)v_p\,] = \frac{1}{a^2}(\dot{y}u_x - \dot{x}u_y)e_3$$

and

$$[\,\widetilde{\xi}, \widetilde{\omega}(p)u_p\,] = \frac{cu_x}{a}e_1 + \frac{cu_y}{a}e_2 - \frac{C}{a^2}(xu_x + yu_y)e_3.$$

Then, we have

$$d\widetilde{\omega}_\xi(p)(u_p, v_p) = 2\frac{mk^2 C}{a^2}(u_y\dot{x} - u_x\dot{y}), \tag{5.5.10}$$

$$\Gamma(p)(v_p, u_p, v_p) \equiv 0, \tag{5.5.11}$$

$$\langle\widetilde{\mathbb{I}}(p)\widetilde{\xi}(p), \widetilde{\Omega}(p)(u_p, v)\rangle = \frac{mk^2 c}{a^2}(\dot{y}u_x - \dot{x}u_y), \tag{5.5.12}$$

$$\langle(D_p\widetilde{\mathbb{I}}(p)u_p)\widetilde{\xi}(p), \widetilde{\xi}(p)\rangle \equiv 0, \tag{5.5.13}$$

$$\langle\widetilde{\mathbb{I}}(p)\widetilde{\xi}(p), [\,\widetilde{\xi}(p), \widetilde{\omega}(p)u_p\,]\rangle \equiv 0, \tag{5.5.14}$$

$$\Xi(p)(u_p, v_p) = mk^2(c - C)(\dot{y}u_x - \dot{x}u_y). \tag{5.5.15}$$

Substituting (5.5.8)-(5.5.15) into (5.5.7), one gets

$$\ddot{x}u_x + \ddot{y}u_y = (-\frac{k^2 C}{a^2 + k^2}\dot{y} + \frac{a^2}{a^2 + k^2}\frac{F_x}{m})u_x + (\frac{k^2 C}{a^2 + k^2}\dot{x} + \frac{a^2}{a^2 + k^2}\frac{F_y}{m})u_x, \tag{5.5.16}$$

where u_x and u_y are arbitrary real numbers, or equivalently,

$$\begin{cases} \ddot{x} = -\dfrac{k^2 C}{a^2 + k^2}\dot{y} + \dfrac{a^2}{a^2 + k^2}\dfrac{F_x}{m} \\ \ddot{y} = \dfrac{k^2 C}{a^2 + k^2}\dot{x} + \dfrac{a^2}{a^2 + k^2}\dfrac{F_y}{m} \end{cases} \tag{5.5.17}$$

If $C = 0$, i.e., the platform does not rotate, Equation (5.5.16) leads to Equation (2.2.43) in the example in Chapter II. If we assume $F_x = F_y = 0$, the solution of (5.5.17) is

$$\begin{cases} x(t) = \dfrac{\dot{y}_0}{\ell}cos(\ell t) + \dfrac{\dot{x}_0}{\ell}sin(\ell t) - \dfrac{\dot{y}_0}{\ell} + x_0 \\ y(t) = -\dfrac{\dot{x}_0}{\ell}cos(\ell t) + \dfrac{\dot{y}_0}{\ell}sin(\ell t) + \dfrac{\dot{x}_0}{\ell} + y_0, \end{cases} \tag{5.5.18}$$

where $\ell = \frac{Ck^2}{a^2 + k^2}$ and (x_0, y_0) and $(\dot{x}_0, \dot{y}_0)$ are initial position and initial velocity of the center of the sphere. It is interesting to note that a trajectory of the center of the sphere is always a circle. But, if we fix a coordinates system, say 0-$X_1 Y_1$, on the

rotating platform, the trajectory of the center of the sphere on this coordinates system, or the trace of the contact point of the sphere on the platform, can be very complicated. Figure 5.5.5 shows some cases for different choices of the parameters a and k. In these examples, we choose $x_0 = y_0 = 0$, $\dot{x}_0 = 1$, $\dot{y}_0 = 0$, $C = 1$.

Remark 5.5.3: Equation (5.5.17) is the same as the one in [32], which, however, was derived from the total dynamics using so-called quasi-coordinates. In addition, this system was not identified as a Chaplygin system in [32]. This is certainly not true if a non-Abelian symmetry is considered, as we have shown above. ∎
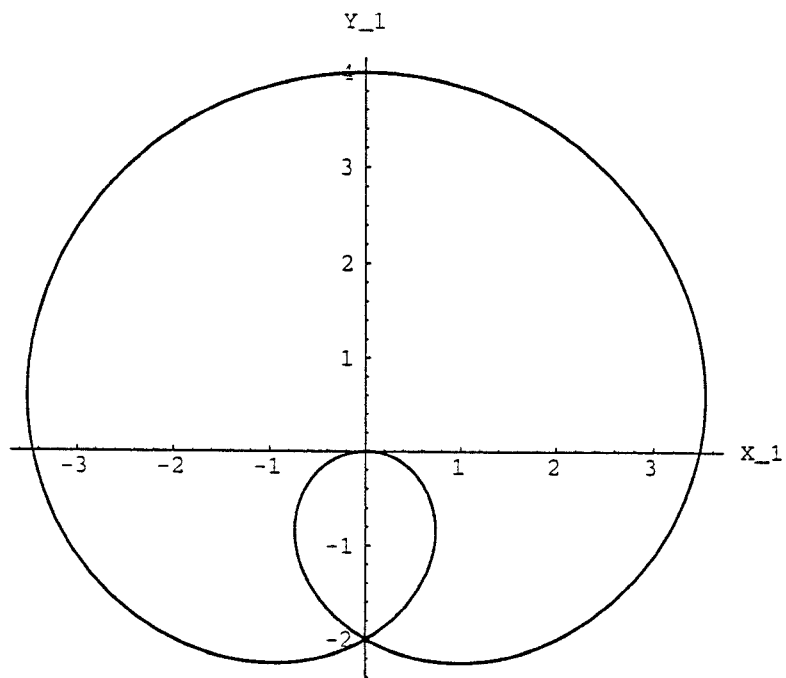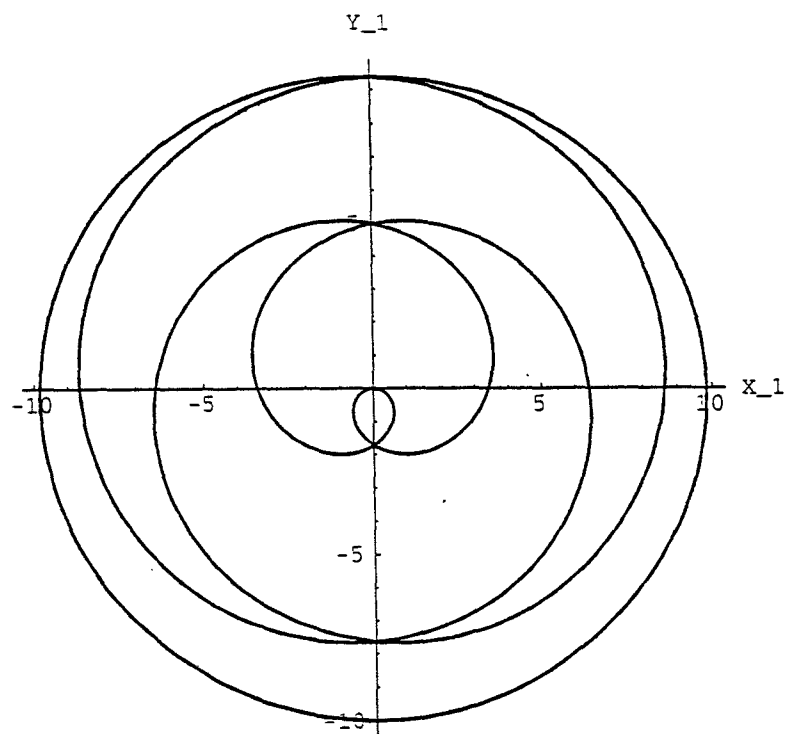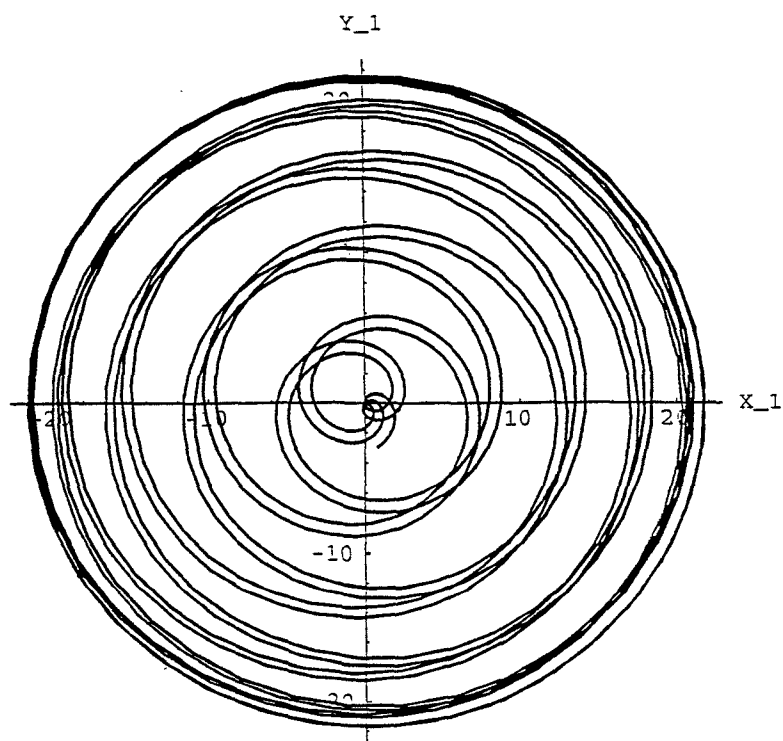
(1). $a = 1, k = 1$
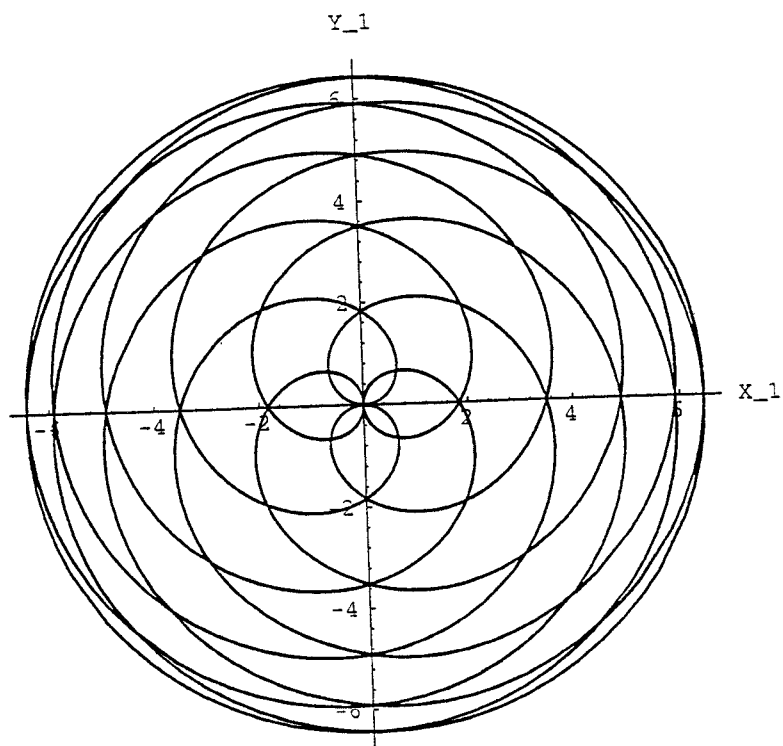
(2). $a = 1.5, k = 1$

Figure 5.5.5 Traces of contact point of sphere on the rotating plane, (1) and (2)

154

(3). $a = 1, k = 0.5$

(4). $a = 3.1415926, k = 1 \ (0 < t < 4.09\pi/l)$

Figure 5.5.5 Traces of contact point of sphere on the rotating plane, (3) and (4)

155

CHAPTER VI

CONCLUSIONS AND FUTURE RESEARCH

Up to now, we have explored various issues concerning with dynamic modeling and kinematic control for constrained mechanical systems with symmetry. The fundamental mathematical tools we applied are group theory, symplectic and Poisson geometry, the theory of reduction in geometric mechanics and the connection theory in differential geometry.

In Chapter II, we reformulate Lagrange-d'Alembert principle with constraints using the classical notion of virtual displacement for constrained systems and a modern treatment of constraints by distributions.

In Chapter III, we rigorously studied the kinematics and dynamics of a particular mechanical system with holonomic constraints: floating four-bar linkage. We have revealed the kinematic and dynamic features of such system in comparison with the systems without geometric constraints.

In Chapter IV, we formulated a kinematic control problem for the system consisting

of a rigid body with two oscillators. From a Hamiltonian viewpoint, we explored the reduction and explicit solvability of related optimal control problems on principal bundles with connection. The necessary conditions for the optimal control problem were determined intrinsically by a perturbation method and a Hamiltonian formulation. By identifying the structure group of the principal bundle as a symmetry group of the system, we were able to use the Poisson reduction procedure to reduce the order of differential equations given by the corresponding necessary conditions. Under suitable hypotheses and approximations, we found that the reduced system possesses additional Abelian symmetry. Applying Poisson reduction again, we obtain a further reduced system and corresponding first integral.

In Chapter V, we turned to the formulation of reduced dynamic equation for nonholonomic Lagrangian systems with symmetry. Under our hypotheses on constraints and exterior force, we showed that the dynamics of a nonholonomic Lagrangian system with non-Abelian symmetry can be reduced to a lower dimensional space determined by the principal fiber bundle. The reduced dynamic equations were formulated explicitly, i.e., without Lagrange multipliers. This formulation generalizes the one for classical Chaplygin systems which possess Abelian symmetry, and the one having non-Abelian symmetry but with linear constraints. In addition, if the mechanical connection of Kummer and Smale is considered, our formulation for nonholonomic Lagrangian systems specializes to the one in Lagrangian reduction discovered recently by Marsden and Scheurle.

We end this chapter by introducing future research relating to the contents of this dissertation.

Concerning the kinematic control problem, more general methods for complete integrability of optimal control on principal bundles with connection will be further investigated. Some methods to study general optimal control problems with nilpotent control algebra may be exploited for this purpose [4,12]. On the application side, the computation of geometric phase and related optimal control problems for more concrete examples, such as a rigid body with flexible attachments, will be studied.

In reduction theory for nonholonomic systems with symmetry we discovered in Chapter V, a gyroscopic type of force appeared in the reduced dynamic equations. Unlike the gyroscopic force coming from the linear term of of the Lagrangian (or internal effect of the dynamics of the system) [50], this force comes from the nonholonomic constraints (or external effect of the environment). Questions about the control of this type of gyroscopic systems and the systems which also include the internal gyroscopic effects will be answered in future research.

There are other types of constrained problems in analytical mechanics, such as systems with Dirac constraints coming from degenerate Lagrangian [10,11] and constrained Hamiltonian mechanical systems, in which the constraints are characterized by the distributions on the cotangent bundle of configurations space [51]. How to deal with reductions for the systems with these constraints is also one of our research interests in the future.

Finally, we note that, in Chapter V, the number of constraint equations in our Constraint Hypothesis is required to equal the dimension of the structure group of given principal bundle. How to relax this condition will also be investigated.

REFERENCES

[1] Abraham, R. & J. E. Marsden, *Foundations of Mechanics* (Second edition), Revised, Enlarged, Reset, Benjamin/Cummings, Reading, 1978.

[2] Arnold, V. I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.

[3] ———, *Dynamical Systems III*, Encyclopoedia of Mathematics III, Springer-Verlag, 1988.

[4] Banks, S. P. & M. K. Yew, "On the Optimal Control of Bilinear Systems and its Relation to Lie Algebra," *Int. J. Control,* 43 (1986), 891–900.

[5] Bates, L., R. Cushman & J. Sniatycki, "Constraints and Reduction," Department of Mathematics and Statistics, Research Report, May, 1992.

[6] Bleecker, D., *Gauge Theory and Variational Principles*, Addison-Wesley Publishing Company, Inc., 1981.

[7] Bloch, A. M. & P. E. Crouch, "On the Dynamics and Control of Nonholonomic Systems on Riemannian Manifolds," *Proceedings of 3rd IFAC Symposium on Nonlinear Control Systems,* , Bordeaux, France (June, 1992), (to appear).

[8] Brockett, R. W., "Control Theory and Singular Reimannian Geometry," in *New Directions in Applied Mathematics*, P. J. Hilton & G. S. Young, eds., Springer-Verlag, Berlin, 1982, 11–27.

[9] Chow, W. L., "Ueber Systeme von Linearen Partiellen Differentialgleichungen Erster Ordnung," *Math. Ann.,* 117 (1940), 98–105.

[10] Dirac, P. A. M., "Generalized Hamiltonian Dynamics," *Canad. J. Math. 2,* (1950).

[11] ———, "Generalized Hamiltonian Dynamics," *Proc. Roy. Soc. London A,* 246 (1958), 326–332.

[12] Faibusovich, L. E., "Explicitly Solvable Non-linear Optimal Control Problems," *Int. J. Control,* 48 (1988), 2507–2526.

[13] Gibson, C. G. & P. E. Newstead, "On the Geometry of the Planar 4-Bar Mechanism," *Acta Applicandae Math.*, 7(1986), 113–135.

[14] Golubitsky, M. & D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory, Volume I*, Springer-Verlag, 1985.

[15] Grashof, F., *Theoretische Maschinenlehre*, Verlag L. Voss, Leipzig, 1883.

[16] Greub, W., S. Halperin & R. Vanstone, *Connections, Curvature, and Cohomology, vol. II*, Academic Press, New York, 1973.

[17] Guillemin, V. & A. Pollack, *Differential Topology*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.

[18] Hrones, J. A. & G. L. Nelson, *Analysis of the four-bar linkage; its application to the synthesis of mechanisms*, The Technology Press of the Massachusetts Institute of Technology, and J. Wiley, & Sons, New York, 1951.

[19] Hunt, K. H., *Kinematic Geometry of Mechanisms*, Clarendon Press, Oxford, 1978.

[20] Koiller, J., "Reduction of Some Classical nonholonomic Systems with Symmetry," *Laboratorio Natonal de Computacao Cientifica Report 17/88* (1988).

[21] Krishnaprasad, P. S., "Eulerian Many-body Problem," in *Dynamics and Control of Multibody Systems*, J.E. Marsden, P.S. Krishnaprasad & J.C. Simo, eds., Contemporary Math. #97, American Mathematical Society, 1989, 187–208.

[22] ———, "Geometric Phases, and Optimal Reconfiguration for Multibody Systems," *Proceedings of American Control Conference*, 3(1990), 2440–2444.

[23] Krishnaprasad, P. S. & R. Yang, "Geometric Phases, Anholonomy, and Optimal Movement," *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, 3(1991), 2185 – 2189.

[24] Kummer, M., "On the Construction of the Reduced Phase Space of a Hamiltonian System with Symmetry," *Indiana University Mathematics Journal*, 30(1981).

[25] Libermann, P. & C-M. Marle, *Symplectic Geometry and Analytical Mechanics*, D. Reidel Publ., Dordrecht., 1987.

[26] Marsden, J. E., R. Montgomery & T. Ratiu, "Reduction, symmetry, and phases in mechanics," in *Memoirs of the AMS,* 1990.

[27] Marsden, J. E. & J. Scheurle, "Lagrangian Reduction and the Double Spherical Pendulum," 1991, preprint.

[28] Marsden, J. E. & A. Weinstein, "Reduction of Symplectic Manifolds with Symmetry," *Reports in Math. Phys.,* 5 (1974), 121–130..

[29] Marsden, J. E. & T. Ratiu, "Reduction of Poisson Manifolds," *Letters in Math. Phys.,* 11 (1986), 161–169.

[30] ———, "Mechanics and Symmetry," University of California, Berkeley, Lecture Notes, 1986.

[31] Montgomery, R., "Isoholonomic Problems and Applications," *Communications in Math. Physics,* 128 (1990), 565–592..

[32] Neimark, J. I. & N. A. Fufaev, *Dynamics of Nonholonomic Systems,* Translations of Mathematical Monographs, American Mathematical Society, Providence, Rhode Island, 1967.

[33] Nomizu, K., *Lie Groups and Differential Geometry,* The Mathematical Society of Japan, 1956.

[34] Oh, Y. G., N. Sreenath, P. S. Krishnaprasad & J. E. Marsden, "The Dynamics of Coupled Planar Rigid Bodies Part II: Bifurcation, Periodic Orbits, and Chaos," *J. Dynamics & Differential Equations,* 1 (1989), 269–298.

[35] Patrick, G. W., *Nonlinear Stability of Coupled Rigid Bodies,* Ph.D. Dissertation, Department of Mathematics, University of California, Berkeley, CA,, 1990.

[36] Paul, B., *Kinematics and Dynamics of Planar Machinery,* Prentice-Hall, Inc., 1979.

[37] ———, "A Reassessment of Grashof's Criterion," *Tran. of the ASME, J. of Mech. Design,* July 101 (1979), 515–518.

[38] Routh, E. J., *Advanced Rigid Dynamics,* Halsted Press, New York, London, 1884.

[39] Rund, H., *The Hamilton-Jacobi Theory in the Calculus of Variations*, D. Van Nostrand Company Ltd, London, 1966.

[40] Saletan, E. J. & A. H. Cromer, "A variational Principle for Nonholonomic Systems," *American J. of Phys.*, 38 (Jul. 1970).

[41] Simo, J. C., T. Posbergh & J. E. Marsden, "Nonlinear Stability of Geometrically Exact Rods by the Energy-Momentum Method," Stanford University, Division of Applied Mechanics, (1989), preprint.

[42] Smale, S., "Topology and Mechanics, I, II," *Invent. Math.*, 10, 11 (1970), 305–331, 45-64.

[43] Sreenath, N., *Modeling and Control of Multibody Systems*, Ph.D. Dissertation, University of Maryland, College Park, 1987, also, Systems Research Center Technical Report *SRC TR87-163*.

[44] Sreenath, N., Y. G. Oh, P. S. Krishnaprasad & J. E. Marsden, "The Dynamics of Coupled Planar Rigid Bodies Part I: Reduction, Equilibria & Stability," in *Dynamics & Stability of Systems 3* #1&2, 1988, 25–49.

[45] Strichartz, R. S., "Sub-Riemannian Geometry," *J. Differential Geometry*, 24 (1986), 221–263.

[46] Vershik, A. M., "Classical and Non-classical Dynamics with Constraints," in *Global Analysis - Studies and Applications I*, Yu. G. Borisovich & Yu. E. Gliklikh, eds. #1108, Springer-Verlag, Berlin Heidelberg New York Tokyo , 1980, 278–301.

[47] Vershik, A. M. & L. D. Faddeev, "Lagrangian Mechanics in Invariant Form," *Sel. Math. Sov.*, 1 (1981), 339–350.

[48] Vershik, A. M. & V. Y. Gershkovich, "Nonholonomic Problems and the Theory of Distributions," *Acta Applicandae Mathematicae*, 12 (1988), 181–209.

[49] Wang, L-S., *Geometry, Dynamics and Control of Coupled Systems*, Ph.D. Dissertation, University of Maryland, College Park, 1990, also, Systems Research Center Thesis Report *Ph.D. 90-5.*

[50] Wang, L-S. & P. S. Krishnaprasad, "Gyroscopic Control and Stabilization," *J. Nonlinear Science,* (1992), to appear.

[51] Weber, R. W., "Hamiltonian Systems with Constraints and their Meaning in Mechanics," *Arch. Rat. Mech. Anal.,* 91 (1986), 309–335.